9-2012

Student Recognition of Visual Affordances: Supporting Use of Physics Simulations in Whole Class and Small Group Settings

A. Lynn Stephens
University of Massachusetts Amherst, lstephens@educ.umass.edu

STUDENT RECOGNITION OF VISUAL AFFORDANCES: SUPPORTING USE OF PHYSICS SIMULATIONS IN WHOLE CLASS AND SMALL GROUP SETTINGS

A Dissertation Presented

by

A. LYNN STEPHENS

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF EDUCATION

September 2012

Education
Mathematics, Science and Learning Technologies

STUDENT RECOGNITION OF VISUAL AFFORDANCES: SUPPORTING USE OF PHYSICS SIMULATIONS IN WHOLE CLASS AND SMALL GROUP SETTINGS

A Dissertation Presented

by

A. LYNN STEPHENS

Approved as to style and content by:

____

John J. Clement, Chair

____

Thomas J. Murray, Member

____

Florence E. Sullivan, Member

____

Clifford E. Konold, Outside Member

____

Christine B. McCormick, Dean
School of Education

DEDICATION

To

Annette Aiken Stephens

who has inspired countless children with her adventurous curiosity and love of learning

and

Wesley Daniel Stephens

who manages to pair an irreverent skepticism with a profound respect for expertise.

ACKNOWLEDGMENTS

I would like to thank my advisor, John Clement, for acting as a mentor in the fullest sense of the word, patiently overseeing years of apprenticeship while being generous with his time, his knowledge, and his professional support. I would also like to thank committee members Florence Sullivan and Clifford Konold for valuable feedback as this project progressed, and especially committee member Thomas Murray for inspiring me to begin this program in the first place and for his continued support and encouragement throughout the process.

My deep thanks to three extraordinarily supportive colleagues: Grant Williams, Norman Price, and Abigail Leibovich. Each of them has offered many hours of feedback, brainstorming, and general enthusiasm. Thank you to Aklilu Tilahun Tadesse, whose example of dedication and determination in doing this work in much more challenging circumstances is a constant inspiration. I extend my appreciation to the high school teachers who took an interest in this work, carved time out from impossibly busy schedules, and contributed many forthright and supportive comments.

And finally, I thank my family, which, jointly, continues to nurture its members with encouragement, appreciation, and financial sacrifices to support our various journeys.

ABSTRACT

STUDENT RECOGNITION OF VISUAL AFFORDANCES: SUPPORTING USE OF PHYSICS SIMULATIONS IN WHOLE CLASS AND SMALL GROUP SETTINGS

SEPTEMBER 2012

A. LYNN STEPHENS, B.S., SUNY EMPIRE STATE COLLEGE

M.A.L.S., SUNY EMPIRE STATE COLLEGE

Ed.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor John J. Clement

The purpose of this study is to investigate student interactions with simulations, and teacher support of those interactions, within naturalistic high school physics classroom settings. This study focuses on data from two lesson sequences that were conducted in several physics classrooms. The lesson sequences were conducted in a whole class discussion format in approximately half of the class sections and in a hands-on-computer small group format in matched class sections. Analysis used a mixed methods approach where: (1) quantitative methods were used to evaluate pre-post data; (2) open coding and selective coding were used for transcript analysis; and (3) comparative case studies were used to consider the quantitative and qualitative data in light of each other and to suggested possible explanations. Although teachers expressed the expectation that the small group students would learn more, no evidence was found in pre-post analysis for an advantage for the small group sections. Instead, a slight trend was observed in favor of the whole class discussion sections, especially for students in the less advanced sections. In seeking to explain these results, qualitative analyses of transcript and videotape data were conducted, revealing that many more episodes of support for interpreting visual elements of the simulations occurred in the whole class setting than in the matched small group discussions; not only teachers, but, at times, students used more visual support moves in the whole class discussion setting. In addition, concepts that had been identified as key were discussed for longer periods of time in the whole class setting than in the matched small group discussions in six of nine matched sets. For one of the lesson sequences, analysis of student work on in-class activity sheets identified no evidence that any of the Honors or College Preparatory students in the small groups had made use in their thinking of the key features of the sophisticated and popular physics simulation they had used, while such evidence was identified in the work of many of the whole class students. Analysis of the whole class discussions revealed a number of creative teaching strategies in use by the teachers that may have helped offset the advantage of hands-on experience with the simulations and animations enjoyed by the small group students. These results suggest that there may exist whole class teaching strategies for promoting at least some of the active thinking and exploration that has been considered to be the strength of small group work, and appear to offer encouragement to teachers who do not have the resources to allow their classes to engage regularly in small group work at the computer. Furthermore, these examples suggest the somewhat surprising possibility that there may be certain instructional situations where there is an advantage to spending at least part of the time with a simulation or animation in a whole class discussion mode.

Keywords: physics education, educational simulations, mental modeling, whole class discussion, small group discussion, science education research, videotape analysis

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF VIDEO CLIPS

CHAPTER I

INTRODUCTION

The following guiding questions motivate the present study.

I. Do students recognize and use affordances and key features of physics simulations and animations? What do we observe teachers doing to support this for students or students doing to support this for each other?

II. What happens when students attempt to reason about key concepts in the simulations? Do they consider causal factors? Do they exhibit conceptual difficulties? If so, to what extent does this get dealt with in discussion?

An important issue that cuts across these questions is that general assumptions about the advantage of small group, hands on work at computers over the use of computers in whole class settings have not been examined. The research was conducted in the context of a larger NSF study on visual modeling strategies in science teaching. The purposes of the larger study provided some constraint on the kinds of data that could be collected and thus on the research questions that could be posed and the methods of analysis that could be used, but the conditions of this larger study also provided a valuable opportunity for conducting the present investigation in the context of both small group and whole class settings.

A. Preliminary Research

In preliminary research, high school physics students were observed routinely missing potential affordances of simulations and failing to attend to key concepts needed in order to understand the material. This work inspired the guiding questions above and suggested constraints and potentially fruitful avenues for the present study.

B. Research Questions

1. Are students able to exhibit gains in conceptual reasoning on pre-post tests from lessons that incorporate certain interactive simulations and animations?
2. To what extent do students and teachers engage in discussion about key concepts while working with the simulations and animations?
3. To what extent do teachers and students respond to conceptual difficulties and misconceptions exhibited during work with the simulations and animations?
4. To what extent do teachers and students support the recognition, use, and interpretation of key visual features of the simulations and animations?
5. Do students recognize and use key visual features of the simulations and animations?

To address these questions, the study employs a mixed methods approach in which quantitative and qualitative methods are used pragmatically in such a way that the resulting combination results in complementary strengths and non-overlapping weaknesses (as recommended by Johnson and Onwuegbuzie, 2004). Quantitative methods are used to evaluate pre-post data, multiple levels of coding are used for transcript analysis, and comparative case studies are used to consider the quantitative and qualitative data in light of each other and to suggested possible explanations.

C. Definitions

By *open coding*, I mean a process of creating codes for transcript segments that are as close to the data and as free from theoretical interpretation as possible, remaining open to unexpected observations and phenomena. By *selective coding*, I mean a process of constraining analysis by coding for selected observation concepts. The theory emerging from the study guided these selections. By *countable or quantifiable coding*, I mean using observation codes that have been refined so that they can produce countable instances or produce some other quantifiable data such as length of time on a given task.

D. Organization of the Chapters

The first part of Chapter II is a review of the literature on the use of visualizations in pedagogy with a focus on model-based science teaching and the use of visual affordances by domain novices. This chapter also includes an overview of the literature on 'scaffolding' as a teaching method, and a discussion of the general theoretical orientation of this work. Chapter III is a report of a pilot study and other preliminary research. Chapter IV outlines the research rationale, design, and method. Chapter V presents the results of pre-post tests administered before and after two short lesson sequences taught as part of normal high school Physics instruction. Chapter VI presents the analysis of selected student work from activity sheets used during the lessons. Chapter VII, the heart of the study, presents the results of the analyses of videotapes of the lessons and contains thick case study descriptions of the classroom discussions. Comparative analyses of the case studies within each matched set of classes examine the qualitative and quantitative data in light of each other. Chapter VIII examines each of the research questions in light of the results and discusses implications for the field.

CHAPTER II

REVIEW OF THE LITERATURE AND THEORETICAL FRAMEWORK

A. Use of Visualizations

1. Importance of Mental Imagery

a. Importance of Visual and Kinesthetic Imagery

Findings from cognitive science (Clement, 1994a, 2004; Hegarty, Kriz, & Cate, 2003; Kozhevnikov, Hegarty, & Mayer, 1999) reinforce the notion of many physicists, e.g., Miller (1986) and Hestenes (1990), that *imagery* is an important form of mental representation in science. Ronald Fink (1989) has defined imagery as *the mental invention or recreation of an experience that in at least some respects resembles the experience of actually perceiving an object or an event*, (but see Hestenes, 1990, for a slightly different take[1]).

Imagery can have components corresponding to any sensory stimuli, as when a subject imagines sensing or applying a force (Reiner and Gilbert, 2000; Gooding, 1992). The work of Clement and others (Clement, 1994; Clement, Zietsman, & Monaghan, 2005) suggests that kinesthetic imagery (mental sensations of how something feels to the touch) can be helpful to creative problem solving in physics. Kinesthetic imagery appears to be associated with physical intuition (Gooding, 1996) and has been used in instruction (Camp et al., 1994; Clement & Steinberg, 2002). In contrast, Sellares and Toussaint (2003) argue that many of the incorrect algorithms recently published in

[1] Hestenes, 1990, distinguishes between *image*, *imagery*, *mental image*, and *mental imagery*. *Image* is a pictorial or diagrammatic representation of information, a *mental image* is a mental representation that is similar to an objective representation, *imagery* the manipulation of an image, and *mental imagery* the mental manipulation of a mental image. This implies that the mental image has some sort of existence apart from the (mental imagery) process; however, I wish to remain agnostic on that point.

computational geometry are due to thinking that contains a strong kinesthetic component. However, their findings suggest that the role of this form of thinking may be more fundamental than previously thought.

b. Importance of Mental Modeling

The ability to generate and evaluate mental models (Johnson-Laird, 1975, 1980) appears to be a crucial aspect of science (Darden, 1991) and of student thinking (Gentner & Gentner, 1983); moreover, it is argued that science textbooks are organized around such models (Giere, 1988). Research continues to indicate the importance of mental modeling in both experts and students (Gentner, 2002; Nersessian, 1995; Nunez-Oviedo, 2004), but Driver (1983) suggests that students often need to be helped to assimilate their prior experience into scientifically accepted models.

c. Importance of Mental Animation

Though some researchers have downplayed the importance of any potentially existing non-propositional aspects of reasoning processes (Forbus & Gentner, 1997; Kintsch, 1986, 1988), Hegarty (1992) hypothesizes that a mechanism involved in subjects' evaluation of their mental models is the use of mental animation to run the models. Hegarty and others have investigated the use of mental animation in problem solving by students (Hegarty, 1992; Clement, Zietsman, & Monaghan, 2005) and experts (Clement, 2006). Some of the mental imagery involved appears to be kinesthetic in nature, as when expert physicists imagine exerting a push or a pull (Clement, 2006; Gooding, 1992).

Hegarty (1992, 2004) believes that students are often induced to animate static diagrams mentally, and this may be a more active learning process than viewing an external animation. Hegarty, et al. (2003), point out that that external visualizations do not always substitute for internal visualizations, citing work from Trickett and Trafton (2002) showing that, even when dynamic visualizations were available to the experts they studied, the experts continued to rely extensively on internal visualization skills and manipulated their internal visualizations more often than they used the computer interface to manipulate the external display. Therefore, Hegarty concludes, we need to foster the development of internal visualization skills. In the present context, this raises a concern that computer simulations and animations may be of limited help unless they foster the development of students' own internal visualizations.

One way of identifying student use of mental imagery is to use *imagery indicators*, observables that plausibly indicate the presence of such imagery. Stephens & Clement (2006, 2010, 2012) use a detailed list of imagery indicators (Monaghan & Clement, 1999) to code videotapes and transcripts of high school physics classes. The model-based lessons incorporated drawings and demonstrations but no animated simulations. The investigators document countable instances of the involvement of kinematic and kinesthetic imagery during videotape episodes where high school science students were generating their own thought experiments. A number of these instances have triangulated evidence from multiple indicators, lending strength to their conclusion that the lessons succeeded in fostering student use of animated imagery, much of it kinesthetic. In related work on student use of extreme case reasoning, Stephens and Clement (2009) find that depictive gestures (gestures that appear to depict an imaginary object, motion, or location in the air in front of the gesturer) and other imagery indicators are associated with many of the student episodes of this form of reasoning; of eight episodes identified, seven are accompanied by depictive gestures that appear to depict either motion or force. This suggests that animated mental imagery was involved. One plausible explanation for these results is that such imagery is important for reasoning and sense making and that a role for extreme cases is to make this kind of imagistic simulation easier, clearer, or more possible for students.

Monaghan & Clement (1999), the study from which Stephens and Clement (2006) drew the list of imagery indicators, identifies evidence that viewing an animated simulation during instruction can result in the use of dynamic mental visualization on problems worked later, away from the simulation. The fact that the subject used hand motions from left-to-right when describing the simulated problem as opposed to up-and-down when referring to the post-test problem indicates that she was transferring the *ability to run the simulation*, not transferring the simulation itself and equating it via analogy. On a less encouraging note, this study also gives evidence of a student mapping the wrong simulations onto problems.

2. Perceptual Affordances of Visual Representations Including Animations and Simulations

a. Early Findings

In the 1980s, cognitive scientists began exploring several issues that would later have impact on thinking about static and animated displays used in pedagogy; these issues include the fact that people possess differing amounts of spatial ability (Just & Carpenter, 1985; Linn & Peterson, 1985). Meanwhile, in the new field of educational research, a few researchers investigated students' ability to visualize scientific processes or entities (Ben-Zvi, Eylon, & Silberstein, 1987; Dwyer, 1987), or to read graphs and diagrams (Seddon & Tariq, 1982; Seddon, Eniaiyeju, & Jusoh, 1984). Clement (1985) and Clement, Mokros, & Schultz (1986) noted several tendencies that appeared to interfere with the ability of adolescent subjects to read or use graphs: the tendency of subjects to view the temporal direction on a graph as a spatial direction; the tendency for subjects to focus on what, to them, is the most salient feature of a graph; a tendency to start all graphs at zero; trouble interpreting interaction between two variables. Clement (1989a) observed that students are often unsure how descriptions or non-numerical aspects of a problem situation map onto a graph. He and others (Kozhevnikov, Kosslyn, & Shephard, 2005) have noted a tendency for novices to treat graphs as pictures. Larkin & Simon (1987) investigated differences between verbal and diagrammatic representations. They noted that verbal representations are essentially linear, which means that when relationships are sought between elements in such a representation, each pairing is separated within the single search dimension. Using two-dimensional diagrammatic representations decreases search time; groups of elements can intersect at one point, can be described by a region, or can be described as contiguous. This can support extremely efficient computational processes, but diagrams can only aid those who can engage in the computational processes necessary to take advantage of them.

b. Short-Term Memory Limited to 4-6 Chunks; Novices Chunk Items, Not Patterns

Larkin, McDermott, Simon, & Simon (1980) report that chess masters can reproduce board layouts from actual games but can’t reproduce random layouts. Most everyone, including the expert, has the capacity to keep in short-term memory only 4-6 chunks. However for the novice, the chunks are individual items while for the expert, the chunks are patterns. The authors believe that a knowledge of large numbers of patterns serve as an index to guide the expert to relevant parts of his or her knowledge store. In the present context, these findings suggest that learning to recognize important patterns in the information in complex visual displays may have a positive affect on students' ability to process and retain information from the displays.

c. Connections Obvious to Teachers May Not Be So to Students; Observations are Influenced by the Theoretical Perspective of Viewer

Driver (1983) has observed that science, formerly taught as a collection of facts, is now taught as a connected system. A problem is that connections obvious to teachers may not be so to students. The theoretical perspective of the viewer (even if naïve) influences observations. Multiple explanations can account for a single set of data; explanations do not spring uniquely from data. She concludes that we do need to present currently accepted scientific theories to students but not in a way that will lead to rigidity. We need to help them assimilate their prior experiences into accepted models.

d. Early Studies with Microcomputer Simulations

Early research on the effects of instruction using microcomputer simulations included studies by Zietsman & Hewson (1986) and Mokros & Tinker (1987). Reiner, Pea, & Shulman (1994) developed and assessed the effectiveness of a set of technology-enhanced teaching and learning activities. These included a dynamic diagram construction kit, hands-on optical tools, and videotape with optical situations and explanatory animations. The activities were used in small groups. The authors believe that conceptual understanding needs to be tested through multiple representational tools because an inability to draw connections between representations can reveal the weaknesses in the conceptualizations. The idea was to foster the development of multiple forms of representation of optical phenomena and the study indicates that this was successful. Many students used multiple representations (verbal, written, and diagrammatic) on the post-test and coordinated the meanings of the representations. The authors also saw progress toward a causal model in students' post-tests. However, students had greater difficulty in using the representational tools and the key (causal) model when presented with less prototypical situations.

e. Dual Coding Theory

Dual coding theory uses a model of cognition that posits two different channels for processing visual and verbal information, resulting in two different memory representations for such information (Paivio, 1986; Mayer, 2003; review by Cook, 2006). Since, in this model, there are separate memory buffers for the two kinds of processes, this theory implies that introducing information into both channels at once can take advantage of the capacity of two buffers rather than just one and can avoid overloading either buffer. However, the information must then be integrated mentally, which adds to the cognitive load of processing the information.

In the research on multi-media tools, this theory has been invoked to predict that animations will work better when paired with oral narration than when paired with written text. A number of studies have applied dual coding theory to the investigation of different combinations of text, graphics, narration, and animation, notably an ongoing series by Richard E. Mayer and colleagues (Mayer, 1997, 2003; Mayer & Anderson, 1991; Mayer & Moreno, 1998, 2002). Indeed, when the material is strictly paced, animation paired with narration appears to work slightly better to promote performance on transfer problems than narration paired with on-screen text (Moreno & Mayer, 1999).

Oddly though, in some studies where printed static shots were taken from the animation and paired with printed text, these static shots plus text produced effects as good as or better than those shown by the animation, whether the animation was paired with text or with narration (Hegarty, Kriz, & Cate, 2003; Narayanan & Hegarty, 2002). This has been interpreted to mean that the material is more important than the medium. However, it would also seem to pose a challenge to dual coding theory, indicating that the pairing of animation with text might not be superior to the pairing of static images with text.

However, Rieber's (1990a, b) studies on fourth graders did seem to show an advantage for animation over static graphics when both of those were paired with text. Thus, it appears that results from experiments using animations, static graphics, and text do not always hold across contexts. It may be that whether one combination of media works better than another, and if so, *which* combination, could be context-dependent (and perhaps strongly so).

Schnotz and Bannert (2003) have put forth an alternate to dual-coding theory. In this integrated model of text and picture comprehension, the processing of text leads to propositional representations and then to mental models, while the processing of pictures leads also to mental models, although through fewer steps. Mental models can have multiple aspects, including auditory, kinesthetic, and haptic. These authors believe that people re-form mental models of a problem in response to the questions asked; therefore, according to this theory, any external visualization materials must support the specific question asked. If such materials don't, even if they are informationally equivalent to materials that do, they could actually hinder the problem solving process by interfering with formation of a mental model that would be more efficient for the task at hand. The experimental results in the study give evidence against dual coding theory and for the integrated theory.

3. Issues with Using Interactive Simulations

a. Recent Findings

Tversky, Morrison, & Betrancourt, 2002, in a review of the literature, indicate that the results of research on the advantage of animated over static graphics is not encouraging. They suggest that animations are often too complex or too fast to be comprehended**.** Lowe (1995, 2003, 2004), for instance, showed that domain novices often miss the important relationships between elements in a weather diagram. Novices instead notice objects or motions that are visually salient but often not important in the causation of the weather changes. Students can also have difficulty interpreting different representations of the same phenomenon (Jones, Jordan, & Stillings, 2001). There has been concern that the effect of animation is not always positive (Mayer & Moreno, 2002; Vikiri, 2002) and that computer animation could conceivably replace students' building of mental imagery. However, others suspect that when that imagery is extremely complex, animation may play a clarifying role (Jones, Jordan, & Stillings, 2001).

b. Importance of Hands on Experience

A recommendation that appears repeatedly in the literature is the following: *Allow students to control the pace of an animation.* This suggestion has been made by a number of investigators (Hegarty, Kriz, & Cate, 2003; Mayer & Chandler, 2001; Schwan & Riempp, 2004; Zahn, Barquero, & Schwan, 2004); these studies found that learners' control over the information flow when learning with computer animations fostered a deeper understanding of the topic. Jones, et al. (2001) suggest giving students power to step through and replay stages in an animation. Hegarty, et al., believe that the user should be able to match the speed of a presentation with the speed of his or her comprehension of the topic. They agree with Tversky, et al. (2002) that animations frequently run too fast for comprehension processes to keep up, and suggest providing VCR-like controls such as Play, Pause, Stop, Fast-Forward, and Reverse. Schwan & Riempp (2004) found that users of non-interactive videos needed about twice the practice time needed by users of interactive videos to learn to tie nautical knots. They suggest the use of an interface similar to that of Apple Computer's QuickTime. Zahn, Barquero, & Schwan (2004) found that students who used functions that allowed them to rewind and forward video sequences, scroll up and down the texts, and go back and forth between text and video learned more. Students who had high frequencies of using the video functions scored higher on the comprehension questions. All of these studies appear to suggest that students who have the opportunity to use a simulation in a hands-on fashion in small groups will do better than students who merely observe and discuss such simulations when projected in front of the class and manipulated by the teacher.

c. Other Recommendations

Recommendations from the literature include the following.

Connect to prior knowledge. Narayanan & Hegarty (2002) recommend that if the targeted users of a multi-modal presentation are not expected already to have the prior knowledge necessary for understanding the conventions of the domain, then connections to prior knowledge should be a part of the presentation.

Have students mentally animate beforehand. Hegarty, Narayanan, & Freitas (2002) found that having subjects first attempt to animate or simulate a system mentally helped them learn more from an animation about the system.

Use of prediction questions may help. Hegarty, Kriz, & Cate (2003) used different combinations of static diagrams, phase diagrams, and animations of a system. Along with each combination, they investigated the effect of adding questions that requested subjects to predict what would happen next to the system. They found that the addition of such questions may have had a small positive effect on comprehension, but this effect did not reach statistical significance in their experiments.

Less may be more. Mayer et al. (1996) found that students who viewed pictures of a process along with brief captions performed significantly better on both recall and transfer problem-solving than did students who viewed the same pictures with a six-hundred word passage that provided more details, whether the passage was presented in a chunk or broken up among the pictures. (The effect size was small).

If a variety of representations is used, explicitly discuss each type with the students (Jones, et al., 2001). To deal with student difficulties interpreting different representations of the same phenomenon, the Molecular Visualization Workshop (Jones, et al., 2001) recommended that the purpose of each type be discussed with the students. Kozma (2003) and Kozma & Russell (1997) examined the effects of multiple representations in chemistry education, and recommend split screen pairings of different representations of the same phenomenon. In addition to studies in chemistry, other studies also suggest that there are drawbacks to using multiple representations (Goldman, 2003).

Provide easy methods to pose questions; have a human tutor available (Jones, et al., 2001). Students exhibit a learning curve with the computer and need help; a learning curve may account for the slower time in the self-paced computer trials of Narayanan and Hegarty (2002). Even if students do not need technical help, the current state of computer interactivity was, at least of the time of the Jones report (Jones et al., 2001), insufficient to provide the same kinds of scaffolding for learning concepts that a human tutor can provide. (It is my impression that this is still the case, at least with most freely available software).

Embed visualization in laboratory work (Jones, et al., 2001, p. 14). The Molecular Visualization Workshop report recommends an iterative process between collection of empirical data and the use of visualization tools to design the next experimental steps.

When selecting simulations, look for those where the linkage between science principles and visual representations is deeply principled. The animation should highlight the causal nature of the phenomena (Hegarty, 1992; Jones, et al., 2001, p. 6).

Some of the above recommendations appear to have little, if any, empirical support, and this investigator does not necessarily endorse them. Nonetheless, they can provide a useful starting point when attempting to identify teacher strategies from lesson plans and classroom videotapes.

B. "Scaffolding" as a Teaching Method

The use of the term *supporting* in the present study is an attempt to generalize the idea of *scaffolding* to contexts in which there is not always a clear distinction between the scaffolder and the scaffolded. Thus, the literature on scaffolding informs but does not constrain the present study.

1. Scaffolding Class Discussion to Deal with Student Conceptual Difficulties

a. History and Description of Scaffolding

Redish (1994) maintains that physics instruction has traditionally appealed to a group of people with a small subset of learning styles and that, because of this, physics teachers are an atypical, self-selected group. He says that undergoing physics training stretches the teachers even further away from the learning style of the "typical" student. Therefore, it can be a challenge to understand what it is that our students need from us in order to understand the concepts we wish them to learn.

The notion of scaffolding grew out of the work of Bruner, a cognitive psychologist (Wood, Bruner, and Ross, 1976) and Vygotsky (1978), a psychologist and social constructivist. Scaffolding is the practice of supporting students to learn concepts that are within their reach with assistance but beyond their reach without it (Hogan & Pressley, 1997). Typically, scaffolding is gradually withdrawn as the student becomes more adept.

One scaffolding strategy is the use of *discrepant events.* Von Glasersfeld (1989) summarized Piaget by saying that cognitive change and learning take place when there is "perturbation" when events do not unfold as the learner's schema would lead him to predict. Discrepant events, then, are those that surprise students because they do not unfold the way the students' current conceptions would lead them to predict. Rea-Ramirez & Nunez-Oviedo (2008) have identified a variant on this strategy that they call *discrepant questioning*, used to help students evaluate intermediate models that have not yet become the target model. It is suggested that these questions initiate dissatisfaction with the child's prior model and motivate him or her to further investigate the concept.

Khan (2008a) describes a scaffolding strategy that she calls *What If?* intended to help students test their mental models. The strategy involves encouraging students to create What If? scenarios where they speculate on and change one or more of the parameters of their model and then observe the effects of the change.

Roth (2001) taught middle school science students using a variety of strategies including having students design machines and present their work to peers, conduct investigations in small groups, and participate in whole class discussions. Roth believes the hands-on experience of using actual pulleys and ropes and drawing diagrams allowed these students, many of whom had problems with academic and social aspects of schooling, to offload part of the cognitive load onto the environment. Doing multimodal presentations allowed a more complex communication than they could have accomplished otherwise, as students articulated and explained their devices and their design choices. Roth believes that learning to model a physical system in the world provides considerable cognitive advantage to students (over working with mental models alone) because it increases the viability of communication and idea development. This raises the question of whether and to what extent computers can allow cognitive offloading, and whether these possibilities differ in whole class use of simulations vs. small group hands-on work at the computer.

b. Scaffolding in Small Group Discussions

Von Glasersfeld (1989), interpreting Piaget, wrote, "(T)he most frequent source of perturbations for the developing cognitive subject is the interaction with others. This, indeed, is the reason why constructivist teachers of science and mathematics have been promoting ‘group learning,’ a practice that lets two or three students discuss approaches to a given problem, with little or no interference from the teacher.” However, studies have reported a variety of issues concerning the effective use of small group discussions in science classes such as the fact that students can exhibit a low level of engagement with tasks (Bennett, et al., 2010).

Jones & Carter (1997), in an analysis of the literature on the effectiveness of cooperative learning, particularly of small group learning, cite Vygotsky’s belief that higher mental learning is the result of social interactions. For instance, students can create metaphors that other students can readily understand. However, the social dynamics of each group must be carefully monitored so that all students have access.

Smith (1996) describes results of Astin (1993) who found that two environmental factors were the most predictive of positive change in college students’ academic development: interaction among students and interaction between faculty and students. These affected more general educational outcomes than any other environmental variables studied, including curriculum content factors. Smith describes two types of cooperative learning groups: Informal Groups that may last only a few minutes and can be used to focus students’ attention on the material, and Base Groups that function long-term and have stable membership, providing students with support, encouragement, and assistance. The author says that faculty need to structure the cooperation; groups must have clear positive interdependence; members must promote each other’s learning; there must be individual accountability along with teamwork skills; and groups need to process, as a group, how effectively members are working together.

A Lumpe & Staver (1995) study of small groups vs. individual study showed an advantage for small groups; students within groups acquired more concepts that were scientifically consistent. Other authors have investigated the workings within groups and teacher strategies for supporting them. Hogan, Nastasi, & Pressley (2000) used discourse analysis to examine the nature and sophistication of peer groups' collaborative scientific reasoning with and without teacher guidance. They found that teachers, though they did not provide direct instruction, acted as catalysts in discussions, prompting students to expand and clarify their thinking. However, peer discussions in the four groups tended to be more generative and exploratory than the interactions with teachers. The authors identified the key acts of participants, both teachers and students, to be working with weak or incomplete ideas until they improved.

During case study analysis, Khan (2003, 2008 a, b) identified small group teaching strategies used by a teacher who employed guided inquiry: a) use of analogies; b) asking students to generate relationships between variables with the use of extreme case reasoning, 'why' questions, and comparisons; c) asking students to compile information using the interactive computer tools to identify variables; d) asking students to work back from the data, predict, design a new test, compare, evaluate; e) asking students to use their evaluations to modify the relationships they had previously generated. Also in case study analysis, Rea-Ramirez (1999) identified a teaching strategy she called 'Explanatory need,' designed to inspire students' need to find explanations for phenomena. This large strategy comprised the use of analogies, discrepant events, hands-on activities, and computer animations and involved discussion that cycled through a series of partial models. This strategy appeared to be successful; all the middle school students in the study successfully constructed mental models of respiration with differences on pre-post tests of more than a standard deviation.

c. Scaffolding in Whole Class Discussions

The following suggestion from von Glasersfeld (1989) has implications for whole class discussion leading: "(C)onstructivist teachers would tend to explore how students see the problem and why their path towards a solution seemed promising to them. This in turn makes it possible to build up a hypothetical model of the student's conceptual network and to adapt instructional activity so that it provides occasions for accommodations that are actually within the student's reach." Minstrell put similar ideas into action in his high school physics classroom with a strategy he calls *reflective discourse* (described in van Zee & Minstrell, 1997a, b), intended to identify and modify students' alternate conceptions. Observations of his classes revealed long silences and long periods of student/student exchanges in whole class discussions. The observer noted that Minstrell greeted student utterances with respect and repeated them with neutral restatements. Discrepant events were introduced only after much classroom discussion. In this study, classroom transcripts were not analyzed by the nature of student response, but by the function that teacher utterances appeared to play within the dynamic of the class discussion.

Hogan & Pressley (1997) suggest that, with prompts and supports, students can become aware not only of *what* they are thinking, but of *how* they are thinking. They suggest ways to expand scaffolding from one-on-one to whole class settings. One method is the circle, where a teacher asks students to direct their remarks to each other and interjects largely to comment on the process rather than to comment on the content of student discourse. The teacher can point out similarities to ways scientists construct knowledge and encourage students to make connections between what they observe and what they already know. In such classrooms, students are observed monitoring their own processes aloud but they have to get used to the fact that they won't always leave class with answers. The teacher supports student thinking rather than compensating for lacks in thinking. The authors observe that the process requires a lot of patience and can be emotionally exhausting for the teacher, and suggest that it be included in the training process for pre-service teachers. They provide practical suggestions for developing scaffolding skills.

Hammer (1995) explores five minutes of transcript of a discussion from his own classroom for what might be seeds of mature science. Students debated whether a ball thrown straight up from a pipe on a moving wagon would fall back into the pipe. There was 20 minutes of discussion between the proposal of the problem and the conduction of the experiment. The teacher/researcher refrained from correcting a student who was maintaining an incorrect position. Only later did he realize that the student had actually articulated a central issue underlying the students' disagreement—and the development of a Newtonian perspective. Although there was little or no evidence in the transcript of traditional content-oriented progress, there was evidence of: a) search for causal factors, b) invocation of prior knowledge, c) construction of thought experiments, d) alternate views considered and addressed, e) key underlying issue identified: is push needed to keep an object moving or to slow it down? f) coherence building. He concluded that the beginnings of science in one student may be very different from beginnings of science in another, so we should not specify, based on a particular model, what one should see in students' work. Rather, the greater the teacher's awareness, the greater the chance of discovering something of value. A teacher can try to support whatever potentially productive elements are found, but must be ready to allow his/her plans for the class to be diverted. Hammer (1997) acknowledges the tension that teachers often feel between the agendas of promoting student inquiry and covering content. In his view, resolution of the tension is not simply a matter of reducing content and welcoming inquiry; rather, it is a matter of responding to students' particular strengths and needs.

In Clement's (2002) study of whole class model construction, he describes different roles teachers can play when they allow student ideas, both correct and incorrect, to be taken seriously in the classroom, though he acknowledges the conflict between content goals and target models. He distinguishes between a *student-directed agenda* and *student-generated ideas* and identifies different pedagogical approaches to foster student creation of new explanatory models. In the 'mosaic approach,' the teacher takes the student ideas, both correct and incorrect, and organizes them: OK, deal with now, deal with later today, deal with after today. However, orchestrating the evolving mosaic mixture can be difficult. Clement speculates that teachers can start from teaching patterns natural to them and then evolve through a competing models pattern to reach their target pedagogical pattern.

Inagaki, Morita, & Hatano (1999) investigate differences in American and Asian teaching styles in mathematics. They found that American teachers tend to give direct feedback to the individual, to "revoice," and to give direct instruction in valid modes of argumentation, while the Japanese teachers encourage students to evaluate each others' arguments, leading students to acquire criteria of evaluation indirectly.

Nunez-Oviedo & Clement (2003) look at model-based teaching strategies that involve model construction at different time scales, including long Macro Cycles that may last up to 2 weeks. Nested within these are intermediate-length teaching cycles such as Model Evolution, which can be of varying lengths. These, in turn, can include smaller, nested teaching cycles such as Model Modification and Model Disconfirmation. All of these cycles, no matter the time scale, can be described as Generate/Evaluate/Modify or GEM cycles. Nunez-Oviedo (2008) looks at a particular kind of intermediate-length cycle, a teaching strategy for guiding whole class discussions that she calls Model Competition. When students in a class suggest ideas that are contradictory to each other, the teacher can use the cyclical Model Competition strategy to support student dissatisfaction with one or more of the ideas. The observed teacher constantly diagnosed the students' ideas and encouraged the students to disconfirm, recombine, restructure, or tune their ideas.

In a series of case study analyses, Williams & Clement (2007, 2009, 2010) identify different levels of strategies intended to foster model construction during whole class discussion, including small-scale dialogic strategies. They analyze how the strategies contribute to cycles of model element construction such as those described in the Nunez-Oviedo studies above and identify a variant Williams calls the OGEM cycle (Observe, Generate, Evaluate, Modify).

Price (2007), a teacher/researcher, documents his own attempt to move from lecture format to a more constructivist teaching practice that uses generative questioning during whole class discussion in order to diagnose current student ideas. He concludes that, just as student ideas may need to change in incremental fashion, our own practices as teachers may need to undergo incremental change in order to move toward more constructivist modes. He suggests small steps that can facilitate the transition.

2. Scaffolding Student Use of Perceptual Affordances of Simulations

Although, as discussed in a previous section, a number of authors have suggested strategies for using simulations, most of these strategies are based on experience, on studies of subjects in tightly controlled laboratory situations, or on theory. It is worth a brief look at the state of empirical research on the effectiveness of these strategies in the classroom, specifically those of the strategies designed to scaffold student use of perceptual affordances.

A number of developers have studied the scaffolding provided by the simulations themselves (review by Cook, 2006) and some have assessed learning from computer-assisted instruction in the classroom (Reiner, Pea, & Shulman, 1995; Raghavan, Sartoris, & Glaser, 1998; Raghavan, Sartoris, & Zimmerman, 2002; Perkins, et al., 2006; Russell & Kozma, 2005). The Reiner, Raghavan, and Perkins studies showed positive effects associated with using the simulations. However, few of these have studied teachers' scaffolding of student use of simulations in the classroom (but see Perkins, et al, 2006; and Price, Leibovitch, & Clement, 2010).

a. Using Whole Class and Small Group Discussions to Scaffold Use of Simulations

Although few, there are some empirical studies that address these issues.

i. Small Group Use

Buckley (2000); Raghavan, Sartoris, & Glaser (1998); and Reiner, Pea, & Shulman (1995) have studied the effectiveness of instructional simulations when part or all of the use was in small groups or by individual students, where 'effectiveness' referred to increased understanding as indicated by student work, student interviews, or student self-reporting on surveys. Williams, Linn, Ammon, & Gearhart, (2004) studied a single teacher and analyzed the kinds of questions she asked and time spent on different kinds of teaching strategies over two years of experience with the Web Based Inquiry Environment (WISE). Other than the Williams study, there do not appear to be many studies that address the question of how to provide instructional guidance for simulations and animations when these are used in small group discussions.

ii. Whole Class Use

Perkins, et al. (2006); Price, Leibovitch, & Clement (2010); and Raghavan, Sartoris, & Glaser (1998) have studied the effectiveness of instructional simulations and animations when at least part of the use was in whole class discussion, where 'effectiveness' referred to usability, interpretation, and learning issues as assessed through student interviews, and in the case of Price, through pre-post tests. Other than these studies, there do not appear to be many studies that address the question of how to provide instructional guidance for simulations in a whole class setting. The Perkins and Raghavan studies do include suggestions to use simulations in this way, but other than Price, there appear to be few, if any, studies that analyze whole class discussion-fostering strategies to support the use of simulations. Jones, et al. (2001) believe we know very little about how to use animation effectively in instruction. Principles suggested by theory and by laboratory work with simulations would appear to need further validation in science classroom contexts (Cook, 2006), and may well have to be modified to be usable by teachers employing available simulations in full class situations.

The only prior study I have found that compared the use of simulations in whole class to the use in small group formats is one by Smetana and Bell (2009), which compared the use of computer simulations in two high school chemistry classrooms taught by a single teacher. They found no significant difference in pre-post gains of the two groups. However, videotape analysis revealed more frequent and meaningful teacher-student interactions and also more frequent highly collaborative talk in the whole-class group. They also note that the whole-class setting can involve topics that extend beyond the pre-planned questions of the lesson. Smetana and Bell suggest that future research involving more varied populations and additional teachers and classrooms is needed.

C. General Theoretical Orientation (Theoretical Framework)

The present study is of classrooms engaged in model-based learning in science (Campbell, 1920; Hestenes, 1987; Clement, 1989). My epistemological stance can best be described as constructivist (Driver & Bell, 1986; von Glasersfeld, 1989), and my pedagogical theory as guided inquiry-oriented (Bell, Smetana, & Binns, 2005; Hammer, 1995, 1997; Herron, 1971) and model-based, especially influenced by the writings of Hestenes (1990, 1996) and my own experiences tutoring and learning physics and mathematics. Findings from social constructivism (Hogan & Pressley, 1997) have led to a belief that classroom discussion that includes student-student exchanges can be an important and helpful component of model-based learning.

1. Model-Based Teaching and Learning: Some Terminology

In *model-based teaching and learning*, a primary goal of instruction is a target model, a desired knowledge state for the student (Clement, 2000b; Harrison & Treagust, 2000; Hestenes, 1987, 1996; Mayer, 1989). This can be contrasted with instruction that focuses primarily on the accumulation of facts and/or practice of procedures. Although mathematical models are taught in the physics courses to be observed in this study, the two lesson sequences of interest focus on *visualizable models* (Clement & Steinberg, 2002; Hegarty, 1992; Hammer, 1995; Hestenes, 1996; Reiser, et al., 2003; Smith, et al., 1997; White and Frederiksen, 2000).

Explanatory models (Campbell, 1920; Clement, 1989, 2000a; Vosniadou, 2002) are scientific models that do not merely represent patterns in observed data (such as PV=KT) but are conceptual inventions that involve invisible aspects that provide explanatory power (such as the pressure equation expressed in terms of numbers of molecules) (Campbell, 1920). Scientific explanatory models include waves, fields, and black holes (Clement, 2000a). An example of an instructional explanatory model is a model of matter as atoms connected by spring-like bonds. In the two lesson sequences to be observed, the simulations explicitly represent visual aspects of the target models, but they only implicitly represent explanatory aspects of the models, as via dynamic relationships between visual elements.

Examples of curricula that have been developed to promote mental modeling are *Energy in the Human Body* (Rea-Ramirez, Nunez-Oviedo, Clement, & Else, 2004) which promotes middle school student development of dynamic mental models of respiration; *CASTLE* (Steinberg & Wainwright, 1993) for middle and high school physical science, which uses air pressure as analogous to voltage differences in electric circuits; *Preconceptions in Mechanics: Lessons Dealing With Conceptual Difficulties* (Camp, et al., 1994, 2010), twelve units focused on specific student misconceptions that have been shown to be persistent in the face of instruction; *Minds on Physics* (Leonard, Dufresne, Gerace, & Mestre, 1999), a complete curriculum for introductory undergraduate physics; and physics curricula that have grown out of the *Modeling Workshop Project* at Arizona State University (Hestenes, 1996; Wells, Hestenes, & Swackhamer, 1995).

A difficult challenge in the implementation of modeling curricula such as these is to keep abreast of the evolving models of multiple students as they participate in large class and small group discussion—daunting even to the most experienced teacher. In fact, Hestenes (1996) states that the most critical element in the successful implementation of the modeling method in the classroom is the skill of the teacher in managing classroom discourse.

CHAPTER III

PRELIMINARY RESEARCH

Part A of this chapter considers several preliminary, exploratory interviews concerning Projectile Motion animations and the implications these have for the dissertation study. Part B discusses results of preliminary analysis of four transcripts of a lesson on Gravitational Potential Energy.

A. Comments from Exploratory Interviews on Projectile Motion

This section considers possible implications arising from comments made during several interviews conducted in the exploratory phase of a large NSF study. These comments were identified during a preliminary phase of analysis, in which the author read the transcripts and checked them against her observation notes. Although conclusions will not be drawn from the comments here, they can help identify potentially fruitful avenues for research, and by doing so, suggest some ways to focus and constrain data collection in a full study.

After conducting a lesson sequence in projectile motion, which the author observed, a high school physics teacher expressed the desire for a projectile motion simulation with characteristics different from those she had been able to find available on the Internet. She liked a simulation she had used early in the sequence that allowed students to investigate the dependence of the range of the projectile on its launch angle (freely available at http://galileoandeinstein.physics.virginia.edu/more_stuff/Applets/ProjectileMotion/jarapplet.html and referred to herein as the "Galileo Simulation," see Figure 24). However, she hoped that by the end of the lesson sequence, students would have constructed a target model of the motion that included the independence of horizontal and vertical components of the velocity, and she did not feel that this simulation, or any other that she had found, adequately addressed this aspect of the motion. In place of an animation, she had used a stop-motion photograph of a projectile in flight. The photograph overlaid a series of snapshots taken at equal time intervals to create a photographic motion map of the kind used by early researchers in motion studies. However, the teacher was not sure how many of her students were able to utilize the visual information in the photograph to conceptualize the independence of the horizontal and vertical components of the motion. She and the author sketched several ideas and the author used equations of motion to create several simulations in Graphing Calculator (Copyright 2007, Pacific Tech, http://www.nucalc.com/), then saved them as short Apple QuickTime animations (Video Clips 1 - 3).

The following semester, the teacher taught the same lesson sequence to three matched classes. By "matched" was meant that students in the classes were comparable in terms of age, they had demonstrated similar levels of aptitude for the content of the course as evidenced by their prior work in the course, and the classes had provided similar levels of preparedness for the lesson. Curious about which version of the animation would work better in place of the stop-motion photograph, the teacher decided to teach one class using the photograph, another class using one version of the new animations (Video Clip 1: Vectors Animation), and a third class using another version of the animations (Video Clips 2 and 3: Lines Animations). All of the classes had seen the Galileo Simulation (Figure 24) earlier in the lesson sequence.

The teacher predicted that the students would like the Lines Animations better than the Vectors Animation because in her experience, this population was not comfortable with vectors if they had studied them at all. The author observed the three classes.

In a follow-up interview after the class, the teacher reported that she thought the Vectors Animation had worked better. Two excerpts from the author's interview notes are below. The teacher is referred to as T and the author/researcher as R.

Excerpt 1
T: I was expecting an 'aha' effect from the lines but—
R: Wish we could get more feedback *(from the students).*

Excerpt 2
T: If I had to do it again tomorrow, I would use the vectors (*animation*), I wouldn't use the lines (*animation*) because they focus too much on distance and not enough on velocity. Another way you could do it is leave ghost images up when it is looping (*referring to the disappearing dots in the Galileo Simulation*).
R: I had the feeling that the vector one gave you more of a lead-in to talk about components.
T: Yes, it did! I had the feeling they were getting the components. We keep coming back to them.
R: If they are ready for it, the vector one is more informationally rich.
T: It's a question of how much trig they would have done when they come in.

The teacher appeared to believe that the Vectors Animation had worked better, but I wondered whether this was so because she had found the vectors easier to talk about. Feeling the need for more feedback from the students, I arranged follow-up interviews with 3 students from the class with the photograph, 3 students from the class with the Vectors Animation, and 2 students from the class with Lines Animations I and II. During the course of each interview, each student was given an opportunity to see all of the animations and respond to them. An excerpt from one of the interviews is below. (Sample questions used to guide these interviews are in Appendix A.) S1 had been in the

Control lesson sequence, which had seen the Galileo Simulation and the stop-action photograph but no animations.

Several minutes into the interview, I brought up Lines Animation I, which highlights the *vertical* component of velocity. S1 looked at the animation, hands on controls. I asked S1 what information she was getting from the display.

> S1: … it's, like, getting bigger, then it reaches its maximum height, then it goes back down and accelerates…

A few moments later, she responded to Lines Animation II, which highlights the horizontal component.

> S1: Well, this is just the same as the other one (*Lines Animation I),* except it's showing where the vertical component is at a given time.

Although this was not correct, I decided to respond neutrally.

> R: OK, and what would you say about that? What information do you get from the red lines?
> S1: Well, like, on the other one, the space between the lines was different sizes. Because, um, between the different intervals it's traveling at different rates. But then here, um, it's not measuring the horizontal velocity, it's measuring the vertical velocity. So that's why they're all the same distance apart.

I had observed some students in these classes express uncertainty over the definitions of "horizontal" and "vertical."

> R: So-- they're the same because the vertical velocity is-- or, OK, maybe you can point and show me which way the velocity is going?
> S1: What do you mean?
> R: Like, you mean the velocity headed in this direction [*gesturing vertically*] or the velocity headed in this direction [*gesturing horizontally*]?
> S1: Yeah, the Y-direction [*gesturing vertically*].

It seemed important to probe to make sure the student actually meant that she thought that the velocity in the vertical direction was constant and that acceleration was occurring in the horizontal direction. After several attempts at probes, which resulted in ambiguous or incomplete answers from the student, the following exchange occurred.

> R: So here [*pointing to Lines Animation I*], you said the velocity was changing. And is that the velocity in this direction [*gesturing vertically*], the Y-component, the vertical velocity that's changing?
> S1: No, this is the horizontal components.
> R: Of the velocity.
> S1: Yeah.

Although, from the full interview, it is not clear that the student actually believed that acceleration was occurring in the horizontal rather than the vertical direction, or indeed that she had any stable understanding of the direction of acceleration, it *is* clear that she interpreted the horizontal lines in Lines Animation I to be indicating something about the horizontal component of motion and the vertical lines in Lines Animation II to be indicating something about the vertical component of motion. Although the teacher and I did not anticipate this misinterpretation, in retrospect it is explainable; the directions of the lines are the most salient visual aspects of these two displays.

Quotations from other interviews (S6: "mm, well it's horizontal movement, so the lines are horizontal," etc.) support the theory that many of these students were more confused by the Lines Animations than they were by the Vectors Animation, and also that many of them were uncertain of the direction of acceleration. Another issue that emerged from the interviews was that many of the students were not clear what was causing the acceleration in the vertical component, and that even fewer of them understood why the velocity in the horizontal component was constant.

1. Implications and Suggestions for Future Research

The purpose of these exploratory interviews was to suggest directions and constraints for future research. There were two outcomes:

1. Transcripts and observation notes helped inform the design of materials to be used by several teachers in a new Projectile Motion lesson sequence.

2. Transcript and observation notes suggested several areas of investigation likely to be fruitful in a larger study, each motivated by a question:

    - Can students correctly identify vertical and horizontal representations in a simulation and map them to the phenomena they represent?

    - Can students correctly identify the changing length of the velocity vector as indicator of acceleration?

    - What is the length of class time spent on the reasons for acceleration in the vertical direction?

    - What is the length of class time spent on the reasons for lack of acceleration in the horizontal direction?

By suggesting fruitful avenues for research, these questions also suggest possible ways to constrain data selection and the analysis of that data in a larger study.

B. Pilot Study: Preliminary Analysis of Four Gravitational Potential Energy Discussions

Although it has been recommended that computer simulations be used with students working hands-on at computers (Jones, Jordan, & Stillings, 2001) and many online educational simulations appear to be designed with that use in mind, in my experience, many teachers have not had ready access to the number of computer stations required for small group hands-on work. However, when simulations and animations are used in a whole class format—for example, projected in front of the class onto a whiteboard—teaching can all too easily devolve into a show-and-tell format, and students may not engage in the kind of active learning that most hands-on activities appear designed to encourage. From a constructivist standpoint, I am interested in what comparisons can be made between the learning taking place during use of interactive simulations in Whole Class situations and that taking place in Small Group situations.

Considering the fact that the hands-on activity afforded by small group work would appear to offer students a more active learning experience, and considering that the teachers in the study stated they prefer to allow students to work with simulations in small groups and feel experienced teaching in that format, it might be expected that the small group format would work better for them. On the other hand, one study has reported a variety of concerns regarding the effective use of small group discussions in science classes, such as the fact that students can exhibit a low level of engagement with tasks (Bennett, Hogarth, Lubben, Campbell, & Robinson, 2010) and another small study reported no significant difference in outcomes after use of simulations in whole class and small group formats (Smetana & Bell, 2009).

This preliminary case study analysis asks:

- *What teaching moves do we observe in small group and whole class work with simulations?*
    - Specifically, what strategies do we observe teachers using to guide discussions to promote conceptual understanding and the development of mental models?
    - What differences do we observe between teacher moves in small group and whole class work with simulations?

This preliminary analysis uses case study comparisons to explore in detail what happened in response to (and in one case, in anticipation of) a single prompting question on an activity sheet used in four classrooms comprising two sets of matched classes (described in more detail in the next chapter). I list some major teacher moves used in the two conditions, drawing from teacher interviews and observation notes. This analysis suggests some unexpected avenues for further investigation in the larger study.

1. Method

a. Participants

68 junior and senior high school students (11th and 12th grades) participated in four physics class sections taught by two teachers in a school in a small, upper-middle class suburban town. The classroom observations were conducted as part of a larger, 3-year study.[2] Participation for each student was voluntary with provisions made for any student who wished to remain off camera. However, almost all of the students in these classrooms elected to participate.

b. Materials and Procedure

A short lesson sequence on gravitational potential energy was taught to matched sets of class sections using lesson plans that incorporated online simulations. For each matched set, the teachers used the same simulation, activity sheet, and other materials in the two conditions but varied the way in which the simulations were used. In the whole class condition, the teachers used a single computer projected onto a screen in front of the class and guided a whole class discussion as students worked through the activity sheet. In the small group condition, multiple computer stations were available with 2-4 students to a computer and the students were allowed to engage in hands-on exploration guided by the activity sheet. In both conditions, the teachers began by introducing the computer

[2] The three years were referred to as Year 0, Year 1, and Year 2. Pilot studies and exploratory interviews were conducted during Year 0 while the classes included in the main study were conducted during Years 1 and 2. Four of the Year 1 classes underwent the preliminary analysis discussed here. The full analysis will be discussed in Chapter VII.

activity in a whole class format, though the extensiveness of this introduction varied. In both conditions, the teachers were available for questions the entire time the simulation was in use. Other than the constraints provided by the teaching materials, the technological set-up, and the data-collection needs of the study, the teachers were free to conduct their classes as they saw fit and were encouraged to use the best teaching strategies they could devise for each situation. The same activity sheets and other materials (manipulatives; prediction sheets asking students to predict what would happen in a system) were used in the two conditions, and control for time on task was implemented by using the same number of class periods to cover the material. The lesson plans, activity sheets, and prediction sheets were developed by the teachers and reviewed by the research team. The teachers selected the simulation ahead of time from online sources.

One teacher taught two lower-level College Preparatory physics class sections and another taught two mid-level Honors physics class sections. Each teacher taught the sequence to one section in small group format and the other in whole class format. The simulation was *Energy Skate Park* http://PhET.colorado.edu. The students used an activity sheet (Appendix B) to guide discussion and to write their ideas, and were administered pre/post tests, but these written data were not analyzed as part of the preliminary study. Preliminary videotape and transcript analyses were conducted.

2. Preliminary Qualitative Analysis and Discussion

It was the author's impression that at times the discussion in the whole class condition was richer than that in the small groups. However, I was not at all sure analysis of the transcripts would reflect this; as might be expected, in whole class discussion students were occasionally seen with their heads on their desks in what appeared to be a "couch potato" reaction. In this preliminary study, I began analysis by examining what happened in the four class sections in response to one of the questions on the activity sheet. I then broadened the scope slightly by looking for any discussion of the topic raised by that question wherever it might appear in the transcripts of the discussions. The intention was to use these matched discussion segments to begin to investigate what aspects of the discussion appeared similar and what appeared different in the small group and whole class settings.

The gravitational potential energy lessons were centered on "Energy Skate Park," a simulation from the PhET project at the University of Colorado (http://phet.colorado.edu/index.php). The simulation has sections of track that can be rearranged and shaped, and several skaters with different masses that can skate on the track. It has a variety of visual tools to help students make sense of the animated imagery and to focus on the abstract quantities under discussion: pie charts, bar graphs, a movable reference line to indicate the height chosen as the zero for gravitational potential energy, a ruler, animated line graphs. In addition, there is an option to have the skater leave behind a trail of dots, each of which can be clicked to obtain a read-out of quantities associated with the skater at that point in the path. The user can change the value of gravity by moving the skater and track to different planets or into space. Friction can be turned off or on and there are thrusters that can apply forces when in space.

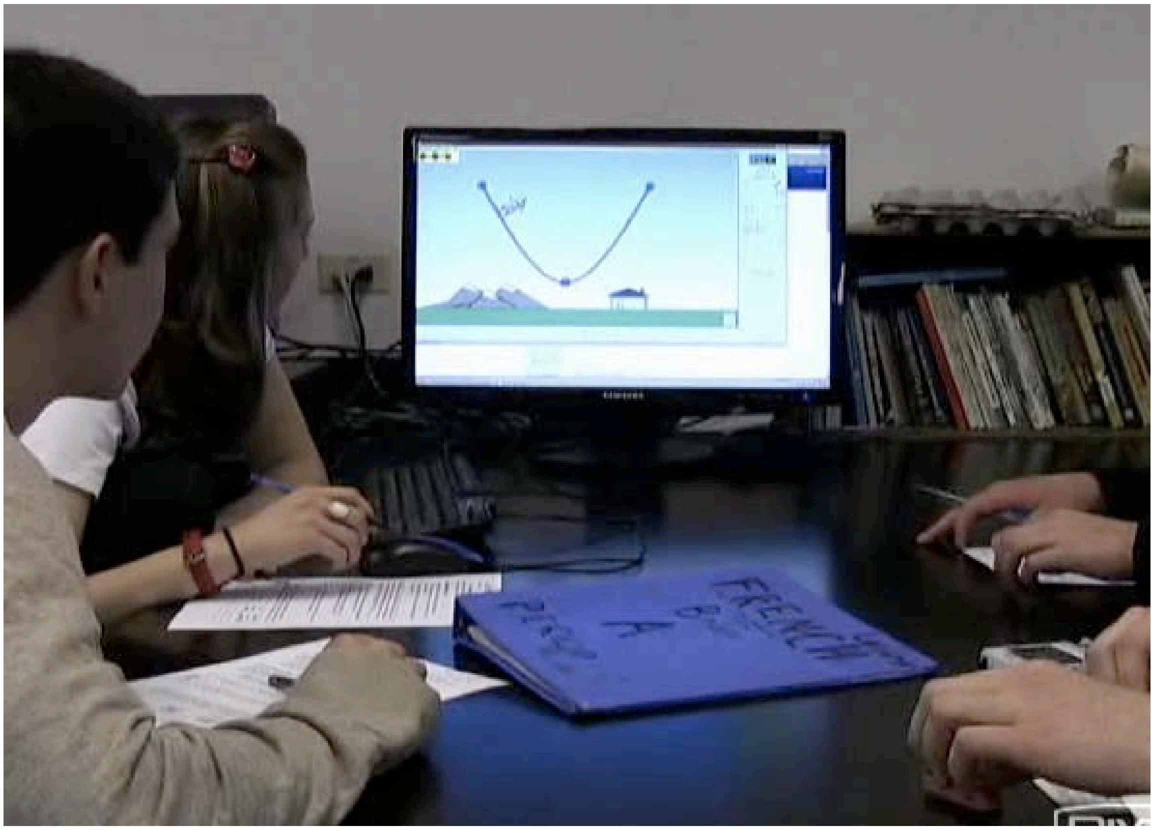


Figure 1: Small group working with *Energy Skate Park*, a PhET simulation

Much of the lesson focused on the parabolic-shaped track shown in Figure 1. The teacher and students referred to this track configuration as a “half-pipe,” though an actual half-pipe does not have this geometry. Objectives of the lesson were for students to begin to understand how potential and kinetic energy can change into each other, the relationship between gravitational potential and height, the arbitrary nature of the choice of potential energy reference line and how this choice affects the measured values of energy, and the relationship between gravitational potential energy and gravitating mass. The physicist’s idea is that the gravitational potential energy possessed by the skater at the top of the half-pipe is converted to kinetic energy as the skater accelerates toward the bottom of the half-pipe. The total of these energies remains constant unless the skater gains energy from or loses it to the environment. However, the value assigned to the potential energy—and, thus, the value calculated for the total energy—depends on the elevation the user has chosen to be at zero potential.

From past experience, the teachers had identified the idea of an arbitrary zero potential reference height (represented by the movable GPE reference line) as a particular stumbling block for their students, especially at the Honors and College Prep levels. A related conceptual difficulty was the idea of the existence of negative energy, especially negative total energy. One of the questions on the activity sheet (Appendix B) was designed to address this directly; Question 7 asked, "Could the total energy be zero at some position? Explain." The correct answer is yes—for example, in the absence of friction, this happens when the zero potential energy line is located where the skater comes to instantaneous rest at the top of his arc.

Observations during a prior year had indicated that this was a topic that had provoked student questions in both small group and whole class discussion. It was thought that analysis of the discussions on this topic in these four class sections should allow a window onto how this topic was dealt with this year. Teacher A taught the lesson to his sections over a two-day period, while Teacher B elected to teach it to her sections in a single day.

a. Gravitational Potential Energy Year 1 Honors Physics Teacher A

i. Small Group Discussion

There were four students in the group, two on each side of a lab table, with the computer on the table between them. The computer controls were accessible, at least initially, to all four students. The teacher circulated the room throughout the discussion, checking in with groups and answering questions.

Of potential concern was the fact that the back and forth between the students did not develop into a substantial discussion of the concepts. Their total time talking and writing about Question 7 was only about 1 minute and they agreed that the total energy could never equal zero. The topic of a zero value for energy did not arise again for this group and a few minutes later they announced that they were done with the activity sheet, even though they had a full additional day to work with the simulation if they wished. (Some, but not all, of the other small groups in the class continued the activity well into the second day. In fact, some of them used almost as much time on task the second day as the first day. It was up to each group how long they spent, however. After groups completed the activity, many of them explored the simulation in other ways.) This discussion will be analyzed in more depth in the larger study.

ii. Whole Class Discussion

This section was taught on the same two days and used the same materials as the class section above. However, in the Whole Class condition, the teacher did not reach Question 7 until the second day of the lesson sequence. Counting the time on task from the first day, Question 7 was reached almost an hour into the Whole Class lesson sequence as compared with less than ½ hour into the sequence for the Small Group described above. (The two class sections had similar times on task: Whole Class used 67 minutes for the two days while Small Group used 61 minutes. However, they did not necessarily use this time in the same ways.)

The discussion of Question 7 took 2¾ minutes as compared to the 1 minute spent on the topic by the small group. It began similarly: the question was read aloud and a student gave a quick answer in response, to the effect that there would have to be no potential energy and no kinetic energy ("On the ground, not moving"). However, there was a subtle difference in the reading of the question in the whole class—the teacher rephrased the question as soon as he read it, making it more active ("How could we get…?"). The student who answered may have understood the concept of zero point energy better than the speakers in the small group that happened to be on camera in the other class, and that could have helped to facilitate the discussion. However, it was this researcher's impression that, no matter the nature of the student response, in the whole class discussion there was often follow-up from the teacher.

Not only was more of the whole class discussion time spent on the existence of non-positive energy values than in the small group observed, but the *student* input on this topic in this whole class discussion, though the teacher spoke frequently, was still more than the student input on the topic in the small group. This factor may have helped compensate for the lack of hands on opportunity afforded the small group students. If the same pattern were to be observed in other matched sets of class sections, this would suggest one possible direction for further qualitative analysis.

Other teacher contributions to the whole class discussion were: making sure that helpful features of the simulation were used; pointing to critical features of the simulation that may have otherwise been overlooked; and appropriating student-initiated ideas into the discussion to keep it going, even when the student ideas were incorrect.

b. Gravitational Potential Energy Year 1 College Preparatory Physics Teacher B

i. Whole Class Discussion

For the discussion of the matched set of physics class sections taught by Teacher B, the Whole Class condition will be discussed first for narrative reasons. She used the same activity sheet and other materials that Teacher A used. She gained a little time by giving the pre-test on the previous day and by instructing the students to skip Questions 5 and 6 on the Activity sheet. The lesson sequence was taught to the College Preparatory sections several weeks later in the term than it had been to the Honors class sections.

In this class, the question about negative total energy arose before the discussion had reached Question 7. The physicist's idea is that the gravitational potential energy possessed by the skater at the top of the half-pipe is converted to kinetic energy as the skater accelerates to the bottom of the half-pipe. As he or she moves, friction causes some of the kinetic energy to be converted to thermal energy. The total of these energies remains constant unless some new energy is introduced to the skater.

The whole class discussion lasted 45 minutes with a large number of student-student exchanges. Even though the teacher took a fairly strong hand in guiding the discussion, she was willing to take cues from students and to try their suggestions for operating the simulation. Occasionally she challenged the students with a question, "What could I do to maybe make his total energy be not so positive?" One student thought she knew how to get the total energy to zero and called out instructions for manipulating the mouse that the teacher followed, resulting in the zero potential energy line being positioned at the top of the skater's arc. This did produce a total energy of zero, though the salient visual on the computer display was the sight of the kinetic and potential energy bars on the animated bar graph swinging wildly up and down in opposite directions. Eventually, the teacher stilled the skater at the bottom of the half-pipe, where she had placed the zero potential energy line for the moment, and all the energy bars registered zero. But she suggested that this was not the complete answer. Finally, she

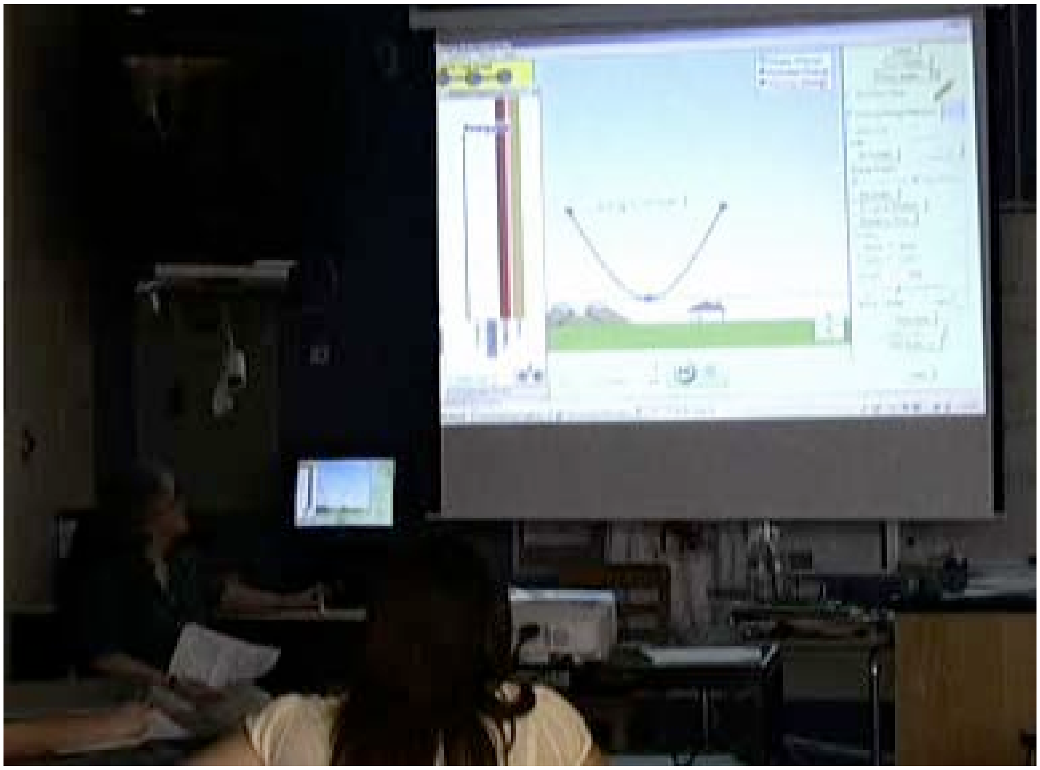


Figure 2: Student Surprise "Wait, he had negative potential energy, what?"

prompted her students to write an answer that was more than a simple yes or no. This rich discussion will be analyzed in greater detail in the larger study.

ii. Small Group Discussion

Teacher B taught a matched Small Group discussion class on the same day. She began with a lengthy introduction to the simulation in the whole class setting before sending the students back to their individual computer stations. As the students worked in their groups, the teacher circulated the room, answering questions and asking them. The small group being videotaped had two students. They reached Question 7 about a half hour into the lesson (comparable to the timing in the Whole Class discussion), 20 minutes after they had arrived at their station and begun the activity sheet.

The transcript segment begins when one of the students read Question 7 and ends when the two students turned to Question 8.

S2: "Could the total energy be zero at the same position?" No, because you don't lose energy. You don't lose or gain energy.
S1: No, because energy is conserved.

(*Students write*.)

This exchange lasted 27 seconds, including writing. This was the total time spent by this group discussing the possibility of a zero value for energy; negative energies were never addressed. Unlike in the whole class discussion, this small group did not use the simulation to explore Question 7; this had appeared to be fairly typical in the small group discussions observed in trial lessons the prior year. One hypothesis is that these students could have been in a "data collection mode," possibly their concept of what laboratory work is supposed to be. Another hypothesis is that, should these students implicitly have held a strong preconception that energy is a quantity akin to a substance and must be positive, the idea of exploring other options or of testing their ideas with the simulation might have been unlikely to occur to them without prompting.

3. Teaching Strategies

In observation notes from the two whole class and two small group discussions that were subjected to preliminary analysis, the following teaching strategies were noted.

Small and large group conditions

- Teacher asks students to predict the answers for some of the questions they will investigate with the simulation to motivate them to think about important conceptual issues and to help them focus their visual attention on important aspects of the simulation.

Small group condition

- Teacher introduces simulation in whole class environment before sending students off to their groups, demonstrates most of the controls and visual features.

- Teacher circulates from group to group, checks in on progress and answers questions.

- Teacher diagnoses what students are and are not getting from simulation, devises one or more prompting questions in response, circulates and asks each group the same question(s).

- When students ask questions, teacher asks them what they can do to test for the answers, "What can you measure to see?" to encourage them to think of their own ways of interacting with the simulation.

- Teacher has quiet activities planned for those who finish with the simulation early (homework, review).

Whole class condition

- Teacher quickly introduces simulation in whole class environment; shows some, but not all, of the controls.

- Teacher invites students to call out suggestions for manipulating the simulation.

- Teacher allows/encourages a student to come up and operate the mouse.

- Teacher pauses simulation and asks students, ""Who will venture a guess about what will happen next?"

- Teacher asks students what they are seeing, points to important but subtle visual features on the screen.

- Teacher repeats selected student comments, adding emphasis.

- Teacher appropriates student-initiated ideas into discussion even when they are partially or wholly incorrect.

- When students ask questions, teacher asks them what they can do to test for the answers, as though students were at the controls themselves. Then he runs their tests for them if practical.

- Teacher waits several seconds after asking a question before moving on, allows silent time while students think or compose their answers.

- After a question on the activity sheet has been discussed in whole class discussion, teacher allows students to talk among themselves as they write their answers.

- Teacher poses question and explicitly invites students to “turn to your neighbor” to discuss it, thereby providing a small-group or partner discussion experience in the midst of the whole class set-up.

- Teacher offers analogies.

- Teacher offers concrete examples.

- Teacher describes the activity in the simulation as though the students were in the world of the simulation. (“If you were there, you would pump your knees in order to go higher.”)

Although some of the strategies listed in the Whole Class category may have been used by the teachers when visiting individual small groups, these strategies were seldom if ever noted in the small group observations.

As can be seen, even though the technology was arranged either for small group work or for whole class discussion, the teachers occasionally found ways to introduce some aspects and possible strengths of whole class to the small group work (whole class discussion before and/or after the activity) and some aspects of small group to the whole class discussion (turn to your neighbor). Both teachers asked frequent questions, especially during whole class discussion, often answering student questions with further questions.

a. Questions Raised

Teacher B expressed surprised when whole class lessons threatened to take longer than small group lessons, as she had expected the small group students to spend more time exploring the simulations in an open, “play” mode. Also, both teachers appeared to underestimate the time they would spend in whole class discussion. They reported finding themselves deviating from the activity sheets more than expected during the discussions because their responses to student questions frequently triggered more student questions, and these, though fascinating, could lead away from the current problem. Though total time on task was consistent across the groups, how that time was distributed appeared to be different.

The hands-on nature of the simulation was designed to afford a rich exploration of the concepts for the small groups, the activity sheet provided a thought-out and detailed guide, and the teacher circulated the room prodding groups and remaining available for questions; however, students did not always appear to notice interesting aspects of the simulations before them and, if they posed questions, frequently did not appear to know how to explore them.

The above observations provide several motivations for a larger study. Further analysis could inquire into such issues as the amount of time spent in discussing causal factors in the two kinds of discussion and how often students used the potential visual affordances of the simulations to assist with this kind of discussion (as opposed to using the simulation to obtain numerical results for assigned problems, for instance). Although the sample is quite small, the preliminary observations raise an interesting question, whether some whole class activities might have the potential to help compensate for a lack of opportunity for hands-on exploration.

4. Conclusion

In the Small Group transcript segments initially examined, I was surprised to find little discussion, occasional misinterpretation of the intended conceptual focus of a question, and a "get and report the data" mindset. In addition, persistent misconceptions may have prevented some students from even considering or examining some issues. Initial examination of Whole Class transcript segments revealed that there appeared to exist teaching strategies for promoting at least some of the active thinking and exploration that has been considered to be the strength of small group work. However, the above case study observations involve a very limited sample. Research on a larger sample into possible differences in the kinds and amount of difficulties students have recognizing and making use of the intended affordances of computer displays, particularly any differences associated with different instructional modes, is, I believe, an important next step.

CHAPTER IV

METHODOLOGY

A. Purpose, Rationale, and Research Questions

Pilot studies suggest that even experienced teachers may underestimate the cognitive difficulties their students face when working with interactive simulations. The purpose of this study is to investigate the issues that arise when students attempt to recognize and use affordances and key features of interactive simulations when they are using them within the classroom setting. My interest is not in what happens in controlled circumstances or when state of the art, proprietary software programs with artificial intelligent tutors are used, but what happens when students attempt to use simulations or animations of the kind that may be freely found on educational websites by—or created by—experienced teachers for use in their high school physics classrooms. When scaffolding appears to be needed to help with perceptual and/or conceptual difficulties, I look for seeds of workable strategies within the context of what is already working well in these classrooms.

Based on the literature reviews of previous studies in related fields and the results of the preliminary research discussed in Chapter III, the following research questions have been identified:

1. Are students able to exhibit gains in conceptual reasoning on pre-post tests from lessons that incorporate certain interactive simulations and animations?

2. To what extent do students and teachers engage in discussion about key concepts while working with the simulations and animations?

3. To what extent do teachers and students respond to conceptual difficulties and misconceptions exhibited during work with the simulations and animations?

4. To what extent do teachers and students support the recognition, use, and interpretation of key visual features of the simulations and animations?

5. Do students recognize and use key visual features of the simulations and animations?

Another important factor that cuts across these five questions is that general assumptions about the advantage of small group, hands on work at computers over whole class use of computers have not been critically examined. Therefore, each of these questions will be investigated in the context of both small group and whole class use, and assumptions about the advantage of one over the other are examined in light of the results.

B. Participants and Setting

This study uses primary data collected by the researcher as part of a larger NSF study on visual modeling strategies in science teaching. The purposes of the larger study provide some constraint on the kinds of data that could be collected and thus on the research questions that can be posed and the methods of analysis that may be used. An important purpose of the larger study is to compare different teaching modes for incorporating physics simulations into classroom activity. The observed lessons were taught in two modes; some used simulations in whole class mode and the others in small group mode. This shapes the data used in the present study but also offers rare opportunities to explore questions about student use of visual affordances within the context of quasi-experimental comparisons.

The population studied comprised the physics students of four teachers at two high schools, one in a suburban college town and the other in a working class community. Teachers were purposefully selected to be those willing to teach model-based lessons, willing to foster discussions in both whole class and small group settings, and willing and able to use computer simulations and animations as part of their lesson plans. Class sections taught by a given teacher were purposefully selected to be included in the data set according to whether they fit criteria to be considered matched sets, as follows. The teacher must have been teaching at least two comparable sections in a given semester and been willing to conduct the lesson sequence in at least one section in a whole class format and in at least one other section in a small group format. Teachers' evaluations and records were relied upon to determine that the sections within a matched set had students comparable in terms of age and demonstrated levels of aptitude for the content of the course as evidenced by their prior work in the course. In addition, the classes in each section must have provided similar levels of preparedness for the lesson, as indicated by the teachers' records of their lesson plans. Finally, the lesson sequence as taught in the two formats must have been similar, as described in Section C: Intervention, and the class sections must have spent similar amounts of time on the lessons and the pre- and post-tests. Thirteen observed lessons were dropped from the study because they did not meet these criteria; nineteen were included.

This researcher observed seventeen of the nineteen lessons and another researcher on the team observed the remaining two lessons. For all classes, the lessons were videotaped, student activity sheets collected, and pre-post tests given. This researcher conducted follow-up interviews with the teachers. Nineteen of the Year 1 Projectile Motion students participated in follow-up tutoring interviews conducted by the researcher. (These were different from the exploratory interviews mentioned in Chapter III.)

C. Intervention

Two topics from the physics curriculum were taught to class sections using lesson plans that incorporated online simulations and/or animations. All lessons included in the data set were taught in matched sets so that quasi-experimental comparisons could be conducted. Once it was determined that class sections were matched, they were assigned to the WC or SG condition for a given lesson sequence according to practical considerations, such as what else was to occur in that classroom that day and how much time there would be before and after the class to rearrange equipment. Class sections within each matched set met in the same room. Because the order in which the sections met rotated each day, the order in which a teacher conducted whole class and small group discussions varied from day to day no matter which section was assigned to which condition. Within each matched set, the teacher conducted the lesson sequence using the same simulation and/or animations, activity sheet, and other materials in the two conditions (Table 1) but varied the way in which the simulations or animations were used. In the whole class condition, each teacher used a single computer projected onto a screen in front of the class and facilitated a whole class discussion as students worked through the activity sheet. In the small group condition, multiple computer stations were used with 2-4 students to a computer and the students were allowed to engage in hands-on exploration and small group discussion guided by the activity sheet. In both lesson sequences and in both conditions, the teacher began by introducing the computer activity to the whole class. In both conditions, the teacher was available for questions the entire time the simulation was in use. Other than the constraints provided by the technological set-up, the pre-designed activity sheets, the lesson mode (whole class or small group) and the data-collection needs of the study, teachers were free to conduct their classes as they saw fit and were encouraged to use the best teaching strategies they could devise for each situation. Control for time on task was implemented by using the same activity sheets and other materials (manipulatives, prediction sheets asking students to predict what would happen in a system) in the two conditions, and the same number of class periods to cover the material. The lesson plans, activity sheets, and prediction sheets were developed by the teachers and reviewed by the research team, which included the author. The pre-post tests were developed jointly by the teachers and research team and consisted of transfer questions that were not directly addressed during instruction; this was to minimize the possibility of the teachers' teaching to the test and also because I wished to measure conceptual rather than rote learning. Sample activity sheets, prediction sheets, and pre-post tests are in Appendices B - G. (The materials varied slightly according to the level of the physics class but in all cases were identical for matched sections. In other words, in a given matched set, students in whole class mode and in small group mode used identical materials.)

The teachers selected the simulations ahead of time from freely available on-line sources. In one of the Projectile Motion lessons, appropriate simulations were lacking so I used Pacific Tech's Graphing Calculator to design simple animations to supplement the online simulation chosen by the teachers. These were saved as QuickTime movies and were uploaded to the school server.

Table 1. List of Lesson Materials

See Appendices B - G for samples of pre-post tests, prediction sheets, and activity sheets.

Gravitational Potential Energy	*Projectile Motion*
activity sheet pre-post test Energy Skate Park simulation: http://PhET.colorado.edu	prediction sheet Galileo Simulation Activity sheet Projectile Animations Activity sheet pre-post test three Projectile Animations (Video Clips 1-3) Galileo Projectile Motion Simulation: http://galileoandeinstein.physics.virginia.edu/more_stuff/Applets/home.html

D. Role of Researcher

This author is a member of the research team that selected the on-line simulations and designed the lesson plans and pre/post tests. I created the three animations, observed most class periods, assisted with videotaping, took observation notes, conducted follow-up interviews with teachers, and interviewed selected students after some of the lessons were conducted. I then analyzed the quantitative and qualitative data, as delimited for this study as described below.

E. Data Sources/Data Collection

Data sources include student work and classroom videotapes and transcripts. Classroom observation notes and interview notes and transcripts were subjected to preliminary analysis and helped provide a context for analysis of the classroom videotapes.

Table 2: Data sources

	Quantitative (countable, statistical analysis on some data)	**Qualitative**
Main sources	Pre-post tests	Classroom transcripts
	Classroom videotapes	Classroom videotapes
	Student activity sheets (selected questions)	

Also used	Classroom observation notes Teacher interview notes Student interview videotapes and transcripts

1. Classes Observed and Videotaped

Videotapes were transcribed with the use of Transana transcription software (Woods & Fassnacht, 2007). It has been relatively difficult to find teachers willing to commit to teaching in different styles in parallel sections. Thus, the study is necessarily opportunistic, gathering data on parallel sections where these were available. Although 57 class periods were observed and videotaped, only the 33 class periods below (19 complete lesson sequences) met the criteria to be described for inclusion in the main study.[3]

a. Gravitational Potential Energy: School 1

'1 SG' and '1 WC' indicate one class section taught in small group format and one in whole class format, respectively.

Year 1	Honors Physics	Teacher A	1 SG, 1 WC
Year 1	College Preparatory Physics	Teacher B	1 SG, 1 WC
Year 1	Advanced Placement Physics	Teacher B	1 SG, 1 WC
Year 2	Advanced Placement Physics	Teacher B	1 SG, 1 WC

This provided four matched sets of class sections, as described in Section B above, N = 150. Teacher A taught this as a two-day sequence while Teacher B taught it as a one-day lesson. Therefore, 10 videotapes were collected for this lesson sequence from the eight class sections above. My intention is not to draw comparisons between different teachers

[3] The three Projectile Motion classes discussed in Chapter III did not fit the criteria, while the four Gravitational Potential Energy lesson sequences discussed in that chapter did.)

but to compare each teacher's small group lesson to the same teacher's whole class lesson of the same matched set. Identical pre-post tests were conducted immediately before and after each lesson sequence.

b. Projectile Motion: Schools 1 and 2

The classes below were held during different semesters than the classes above, and though some of the same students were involved, the sections had been reshuffled somewhat between semesters. The two lesson sequences are analyzed separately; no student is represented more than once in any given analysis.

Year 1	Honors Physics	Teacher A	1 SG, 1 WC
Year 1	Honors Physics	Teacher C	1 SG, 2 WC
Year 2	Honors Physics	Teacher A	1 SG, 1 WC
Year 2	College Preparatory Physics	Teacher A	1 SG, 1 WC
Year 2	Advanced Placement Physics	Teacher B	1 SG, 1 WC

This provided 5 matched sets of class sections, as described in Section B above, N=200.[4] Teacher A taught this as a two-day lesson sequence, Teacher B taught it as a one-day lesson, and Teacher C as a three-day sequence. Therefore, 23 videotapes were collected for this lesson sequence from the eleven sections above. The analysis focuses on the videotapes that show the students working on the Projectile Animations activity sheet and using the Projectile Animations, although all videotapes were viewed and subjected to

[4] Most students were involved in both lesson sequences but many of these lessons did not fit the criteria to be included in the analyses. Also, some students who did not have a complete set of pre-post tests in one sequence did have a complete set in the other sequence. The two analyses use data from 274 individual students, approximately a quarter of whom are represented in both studies. In addition to these lesson sequences, each student participated in many other sequences throughout the year conducted in whole class and small group formats, the customary practice of these teachers.

some analysis, as described below. Identical pre-post tests were conducted immediately before and after each sequence.

F. Research Design: Mixed Methods Approach

1. Rationale for Mixed Methods Approach

The purpose of the present study is to describe new phenomena and generate new descriptions of student difficulties and teaching strategies, and to develop descriptions that have the potential to apply to other classrooms and topics and that can inform theory and practice. The questions being asked have a strong qualitative component; my intention was to learn as much as I could in these settings about what perceptual and conceptual support students need in order to be able to make use of the simulations and animations, and what strategies educators—and by extension, simulation designers—can use to support those needs. However, I also had the valuable opportunity to control some of the many variables in the classrooms in order to create a quasi-experimental design in a naturalistic setting and to collect quantitative data. The results of pre-post tests can reveal different kinds of patterns than those revealed in the qualitative data; they give some indication of what was learned by students who did not speak up in class and provide opportunities to triangulate different kinds of evidence to support claims. An important difference from traditional quantitative studies is that *the quantitative comparisons of pre-post gains are not being used to attempt to project findings to a population outside the study*. Rather, they are being used 1) to add quantitative detail to the individual case studies by determining whether some learning occurred during each lesson sequence and whether there were differences between the gains of matched sections, and 2) to suggest the interesting presence (or unexpected absence) of trends that may be worth investigating in future studies with larger samples.

a. Strauss & Corbin: Grounded Theory

The methodology of the present study has been influenced by grounded theory (Glaser & Strauss, 1967) especially as it was developed by Strauss & Corbin (1998), in that I wished to begin my analysis of the transcript data by setting aside theoretical perspectives as much as possible and allowing theoretical concepts to emerge from the data. Though influenced by grounded theory, my method was not bounded by it, but rather I followed a pragmatic, mixed methods approach as described below. I began with transcript coding that was more open than theory-laden and progressed toward coding that was increasingly theoretically sensitive. I also used *theoretical sampling*, that is, I decided what data to collect or to analyze next according to the present state of my theory generation process.

b. Clement: Analysis of Clinical Interviews

Clement (2000a) has identified a typology of kinds of transcript analysis along a spectrum from more generative-interpretive to more convergent-coded (Table 2).

According to Clement, studies with *generative* purposes, which aim to generate new ideas and observation categories, usually lend themselves to interpretive analysis; they can deal with behaviors that are unfamiliar and for which there is little in the way of existing theory, and they can deal with larger and richer sections of transcript data.

Studies with *convergent* purposes usually lend themselves to coded analysis; they involve attempts to provide reliable, comparable, empirical findings that can be used to determine frequencies, sample means, and at times, experimental comparisons for testing a hypothesis.

Although some have treated generative and convergent studies as a dichotomy, Clement identifies a spectrum of studies that move from the more generative to the more convergent. In Table 3 below, taken from Clement (2000a) and slightly adapted, he divides this spectrum into four levels; the method of Level B is similar to the constant comparative method of Glaser and Strauss (1967) mentioned above.

Table 3: Spectrum of clinical interview approaches from generative to convergent[5]

A)	*Exploratory studies:* Relatively large sections of transcript are explained by a global interpretation that may contain several elements. The analyst formulates an initial description of the subject's mental structures, goals, and processes in order to provide an explanation for the behavior exhibited in the transcript. This involves the construction of new descriptive concepts and relationships on a case-by-case basis. Examples of transcript sections are usually exhibited in reports alongside the analysts' interpretations. In exploratory studies, sensitivity to subtle observations is important; e.g., investigators may make use of facial expressions, gestures, and voice inflections.
B)	*Grounded model construction studies:* Analysts generate descriptions as in Level A above. In addition, some initial observation patterns are identified. Investigators analyze smaller segments of transcripts and begin to separate theoretical concepts (partial theoretical models) from observations. They also begin to connect theoretical models to specific observations that support them, triangulating where possible. A stable context is needed for those observations that will be compared across different subjects and episodes.
C)	*Explicit analysis studies:* Investigators criticize and refine observation patterns and theoretical model elements on the basis of more detailed analyses of cases; articulate more explicit definitions of observation categories (definitions of observation categories should approach independent codeability); code for certain observation categories over a complete section of transcript according to a fixed definition or criterion. If the study has a theoretical component, investigators point to sets of observations in a transcript and explain them by means of a model; articulate more explicit descriptions of theoretical models; and describe explicit triangulated lines of support from observations to theoretical models.
D)	*Independent coder studies*: Analysts refine concepts as in Level C above. In addition, coding of observation patterns is done by independent coders and inter-rater reliabilities are calculated. Note that it is much easier to define rules or

[5] (Clement, 2000a, Table 4, slightly adapted).

guidelines for coding observable patterns in observations than for coding unobservable theoretical model elements. Coding that is restricted in this way still can provide a strong source of support for a constructed model when coded observation patterns are judged by readers to provide evidence for the theoretical model.

Levels A and B are more generative while Levels C and D are more convergent. Such work can be cyclical; work at a convergent level can initiate work at a generative level and vice versa; just as observation patterns can suggest theories, theories can suggest new observation categories.

The design of the present study could be described as an opportunistic or strategic mix of quantitative as well as qualitative methods from Clement's levels A, B, and C in Table 3. (See comment on triangulation at the end of this chapter.)

c. Johnson and Onwuegbuzie: Mixed Methods Research

Johnson and Onwuegbuzie (2004) have developed a typology of mixed methods designs that include both designs in which a qualitative study is followed by a quantitative study or vice versa, and designs that mix qualitative and quantitative approaches within or across the stages of the research process (though one or the other method may predominate in a given stage). They point out that quantitative data may be converted into narrative data that can be analyzed qualitatively, and qualitative data may be converted into numerical codes that can be represented statistically. They quote Charles Sanders Peirce: "Reasoning should not form a chain which is no stronger than its weakest link, but a cable whose fibers may be ever so slender, provided they are sufficiently numerous and intimately connected" (1868, in Menand, 1997, pp. 5-6). Their philosophical underpinning is the pragmatism of Peirce, William James, and John Dewey. The bottom line for Johnson and Onwuegbuzie is that research approaches should be mixed in ways that offer the best opportunities for answering important research questions. They suggest that researchers "collect multiple data using different strategies, approaches, and methods in such a way that the resulting mixture or combination is likely to result in complementary strengths and nonoverlapping weaknesses" (p. 22).

Strengths of mixed methods research that apply to the present study are that words, pictures, & narratives are used to add meaning to numbers; numbers are used to add precision to words, pictures, and narrative; and that this method addresses a broader and more complete range of research questions. The method has been cyclical, moving back and forth between quantitative and qualitative methods, as the findings of one method inform the implementation of others.

The research design had three stages:

Stage 1: Analysis of pre-post data (predominantly quantitative);

Stage 2: Analysis of transcript data, analysis of selected activity sheet data (mixed methods);

Stage 3: Comparative case study analysis (predominantly qualitative).

2. Pre-Post Test Data Analysis

a. Quantitative (Statistical Methods): Short Answer Pre-Post Data

Most of the questions on the pre-post tests were short answer questions. Appropriate statistical methods such as paired samples *t*-tests were used to test for pre-post gains and to compare average gains between whole class and small group conditions. A statistics expert familiar with classroom research and with the present study recommended conducting separate comparisons within each teacher and each topic rather than trying to pool results or to compare quantitative results across teachers or topics; this advice is followed.[6]

b. Quantitative (Quantifiable coding): Explanation Question Pre-Post Data

The pre-post tests used for both topics included one or more explanation questions. The answers to a selected set of these questions were analyzed and coded according to simple rubrics (blind to condition and to whether test was pre or post). The results of this analysis were not subjected to statistical analysis but are presented in tabular form and discussed in connection with the multiple-choice pre-post data.

3. Videotape and Transcript Data Analysis

a. Qualitative (Grounded Theory Development): Selected Videotape and Transcript Data

The videotapes and transcripts for the observed classes were examined using elements of the constant comparative method to progress from noting observations throughout substantial portions of transcript data, to identifying patterns in observations, to defining codes that were used for selective coding of transcripts. The results of this iterative process are presented as a list of refined codes and coding criteria and discussed separately for the two lesson topics.

b. Quantitative (Quantifiable Coding): All Relevant Videotape and Transcript Data

Using the codes developed and honed in the process of grounded theory development, I used Transana transcription software to conduct selective coding across

[6] This practice resulted in nine small comparisons rather than one large one; however, as explained above in the rationale for a mixed methods approach, the intention of this exploratory study was to identify trends *within* the study that could suggest hypotheses, *not* to attempt to project findings rigorously to a population beyond the study.

all relevant transcript sections to produce countable or quantifiable comparisons of instances of phenomena.

4. Activity Sheet Data Analysis

a. Quantitative (Quantifiable Coding): Selected Explanation Question Activity Sheet Data

Answers to selected explanation questions from student activity sheets were analyzed and coded according to simple rubrics (blind to whole class or small group condition). The results of these analyses are presented in tabular form to be triangulated with certain transcript data to provide multiple sources of evidence.

5. Comparative Case Study Analysis

a. Predominantly Qualitative Methods

First, the results of the quantifiable coding of the transcripts are presented in the context of thick case study descriptions for each class. These results are occasionally supplemented with material drawn from observation notes and follow-up interviews of teachers and students in order to create a rich description of what occurred in that class.

Second, for each matched set of classes, the results of quantitative and qualitative analyses above are subjected to comparative case study analysis. The results of quantifiable coding from the videotapes are presented in graphical and tabular form so that observation patterns can be compared numerically across lesson formats within each matched set. These results are triangulated with activity sheet data when possible, and examined in light of the quantitative results from the pre-post tests. All these data are discussed in terms of possible theoretical implications.

G. Data Selection in Relation to the Research Questions

Following Table 4, which pairs research questions and types of evidence, each research question is discussed in terms of the evidence and research methods used to address it.

Table 4. Evidence to address each research question

Research Question	*Research Method*	*Activity*	*Questions addressed*
Q 1	Stage 1 Quantitative (statistical) methods Quantitative (quantifiable coding) methods	Analyze **pre/post tests** Compute results for all short answer questions Code answers for selected explanation questions	What patterns, if any, are revealed by pre-post quantitative results?
Q 2-5	Stage 2 Qualitative methods Countable or quantifiable coding methods Stage 2 Qualitative methods Countable or quantifiable coding methods	Analyze **transcripts** Do open coding of selected transcripts to generate countable constructs Using countable constructs, code all transcript sections relevant to the constructs Analyze **activity sheets** Do open coding of selected answers to generate countable constructs Code answers to selected explanation questions	What observation patterns can be discerned from transcript data? Are the patterns born out in student responses?

Research Question	*Research Method*	*Activity*	*Questions addressed*
Q 1-5	Stage 3 Predominantly qualitative methods	**Comparative case study analysis** Synthesize, illustrate, and illuminate quantitative and mixed-method findings in comparative case studies	Why might these patterns have occurred?

1. Evidence to Address Research Question #1

Are students able to exhibit gains in conceptual reasoning on pre-post tests from lessons that incorporate certain interactive simulations and animations?

a. Quantitative Analysis

All of the pre-post tests used in the study contained two kinds of questions, short answer and explanation questions. The author scored the answers in conjunction with other knowledgeable members of the team, consulting with an expert on scoring decisions. All scoring was done blind to condition and to whether the test was pre or post. All answers were scored using rubrics developed by the author (Appendices C and G) according to whether they were correct, partially correct, or incorrect. During development of the scoring rubrics for answers to the explanation questions, the author consulted regularly with an expert across the cycles of development of the coding definitions and examples.

For Gravitational Potential Energy, the test was refined between Years 1 and 2 to reduce ambiguities in the questions. After dropping ambiguous questions, both versions were scored for 9 short answer questions. In addition, the Year 1 version was scored for two explanation questions and the Year 2 version for one explanation question.

For Projectile Motion, the Honors/College Prep test was refined between Years 1 and 2 to reduce ambiguities. After dropping ambiguous questions, both versions were scored for 8 short answer questions and 2 explanation questions. The AP classes exhibited a pre-test ceiling effect in Year 1 and those results were dropped from the study. They were given a substantially revised, more difficult test in Year 2, which was scored for 6 short answer and 2 explanation questions.

For the short answer questions, *t*-tests were used to test for the significance of average pre-post gains in each class as well as significant differences in average gains due to class format within each matched set of classes. Effect sizes and confidence intervals were computed in order to evaluate the meaningfulness of gains and of observed differences associated with class discussion mode. The results of the explanation questions were not subjected to statistical analysis but are presented in tabular form.

b. Comparative Case Study Analysis

The comparative case study analyses examine patterns observed in the pre-post data in light of observation patterns identified in the transcripts, and then propose hypotheses to explain patterns observed in the pre-post data.

2. Evidence to Address Research Question #2

To what extent do students and teachers engage in discussion about key concepts while working with the simulations and animations?

Research questions #2-5 were addressed by a mixed methods design using three types of transcript analysis techniques: open coding for concept development, countable or quantifiable coding, and comparative case study analysis.

For question #2, analysis of videotapes, observation notes, and follow-up interviews from the preliminary studies led to the identification of several *key concepts* that appeared important to the students in those classes and for which the simulations/animations appeared to offer strong affordances for development. However, as illustrated in Chapter III, the presentation of these concepts appeared to vary widely across the early lessons observed. Also, discussion about these concepts was at times distributed among multiple participants. Therefore, I did not attempt to count specific instances of student reasoning about the key concepts but measured the total time spent in reasoning about the concepts in each whole class or small group discussion.

a. Open Coding

I used open coding of selected videotapes to identify additional key concepts that appeared important to student understanding in these classes. Descriptions of all key concepts were then developed and honed through an iterative process. This eventually allowed stable codes to emerge that could then be applied to a broader selection of transcripts for efficient and quantifiable coding.

b. Quantifiable Coding

Transcripts from the relevant portions of all matched sets of classes were coded for *amount of time* spent in each class on the specific key concept. I look for patterns in length of time spent, especially patterns associated with a difference in discussion mode. For each matched set of classes, these results are presented in charts and tables that follow the individual case studies and precede the comparative case study discussion.

c. Comparative Case Study Analysis

Thick descriptions of examples of such discussion are provided in individual case study narratives. Any patterns observed in the quantifiable codes are discussed and contrasted in case study comparisons.

3. Evidence to Address Research Question #3

To what extent do teachers and students respond to conceptual difficulties and misconceptions exhibited during work with the simulations and animations?

The intention is to identify when student *conceptual* difficulties were being acknowledged and addressed by the discussion and to look for possible patterns in such support. Supporting discussion may be initiated by the teacher or by another student, as when a student with clear understanding of an issue seeks to help another student by asking a supporting question.

a. Quantifiable Coding

Evidence was obtained from classroom videotapes in multiple phases. First, moments in the relevant portions of transcripts from all matched sets of classes were identified where students appeared to experience conceptual difficulties, exhibiting puzzlement, confusion, or frustration about a conceptual issue. The transcripts were then examined to see if and how long the discussion addressed these difficulties. Also noted were moments where teachers or students appeared to be addressing misconceptions, even if the students being addressed did not exhibit any awareness of having conceptual difficulty. I look for patterns in length of time spent on such discussion, especially patterns associated with a difference in discussion mode. For each matched set of classes, these results are presented in charts and tables that follow the individual case studies and precede the comparative case study discussion.

b. Comparative Case Study Analysis

Thick descriptions of examples of such discussion, and examples of teaching moves observed, are provided in individual case study narratives. Any patterns observed in the quantifiable codes are discussed and contrasted in case study comparisons.

4. Evidence to Address Research Question #4

To what extent do teachers and students support the recognition, use, and interpretation of key visual features of the simulations and animations?

The intention is to identify the amount and kinds of support used to address student *perceptual* and other difficulties in making effective use of key visual features that were intended affordances of the simulations and animations. (Determination of which visual features appeared key for these students is discussed under Q5, below.) Either a teacher or student could employ these moves, as when a student sought to help another student by asking a supporting question or pointing out the visual feature. Analysis of videotapes, observation notes, and follow-up interviews from the preliminary studies led to the identification of an initial set of support moves that appeared to have been useful for the students in those classes.

a. Open Coding

Additional moves that appeared intended to support student recognition and use of the key features were looked for during open coding of selected videotapes. Descriptions of these support moves were developed through an iterative process and honed for use in quantifiable (countable) coding.

b. Countable Coding

Transcripts from the relevant portions of all matched sets of classes were coded for instances of support moves. The numbers of teacher and student moves are tallied and compared.

c. Comparative Case Study Analysis

Examples of teaching moves that appear to have been helpful in supporting the recognition and use of key visual features of the simulations are given thick descriptions in the individual case studies. Patterns observed in the countable codes are discussed and contrasted in the case study comparisons. An additional objective is to provide rich descriptions of moves that have the potential to apply beyond the classrooms from which these moves were drawn.

5. Evidence to Address Research Question #5

> *Do students recognize and use key visual features of the simulations and animations?*

Analysis of videotapes, observation notes, and follow-up interviews from the preliminary studies led to the identification of an initial set of visual features that appeared to provide important support for the students in the preliminary studies as they attempted to reason about the key concepts.

a. Open Coding: Videotapes

Open coding of selected videotapes identified additional visual features and several visual relationships that appeared important in supporting student reasoning about the key concepts. Descriptions of these features and relationships were developed and gradually honed through an extensive iterative process. This honed set of features and relationships was used in countable coding.

b. Countable Codes: Videotapes (Gravitational Potential Energy)

Two key features in the Energy Skate Park simulation were identified as especially important in providing support to students as they attempted to reason about the key concepts. Videotapes from the relevant portions of all matched sets of Gravitational Potential Energy classes were coded for student use of these features. Teacher use of the features was also coded for purposes of comparison. The numbers of teacher and student episodes are tallied and compared.

Several key features and relationships in the Projectile Motion Animations were identified as important in providing support to students as they attempted to reason about the key concepts. Because these animated features could not be manipulated, it was difficult to determine via videotape analysis when students were using them. Therefore, estimates of student use during the Projectile Motion lessons are obtained exclusively from activity sheet coding, as described below.

c. Open Coding: Activity Sheets

Once the lists of key features for each lesson sequence were developed and honed via analyses of selected videotapes, open coding was used on a stratified sample of *student activity sheets* from each lesson sequence to identify evidence in student writing and drawing for use of the features. This coding was done blind to condition (whole class or small group). Activity sheet questions that appeared capable of yielding information about this question were identified (four questions from the Projectile Motion sheet and one or two questions from each Gravitational Potential Energy sheet). Coding criteria for the written and drawn portions of student answers were developed by the author in an iterative process and honed for use in countable coding. During development of the scoring rubrics, the author consulted regularly with an expert. After development of the rubric, the expert and the author each used it to score a sample of the answers; the resulting scores were in agreement.

d. Countable Codes: Activity Sheets

The author used the rubric to code all activity sheets from the matched sets of classes for student written and drawn responses to the selected questions; this was done blind to condition. The results were tallied by the author and are presented in Chapter VI.

Unlike the results of videotape coding, these results are drawn from all students in the classes—including those who were in small groups not on camera. Therefore, they allow a different kind of comparison between the discussion formats than videotape analysis allows. The Projectile Motion activity sheet questions were designed with this research question in mind. For the Gravitational Potential Energy activity sheets, which were not designed with this question in mind, the coding results will be triangulated with the results of the videotape analysis described above.

e. Comparative Case Study Analysis

For the Gravitational Potential Energy classes, examples of student and teacher use of the features identified during videotape analysis are described in the individual case studies. Patterns observed as a result of this analysis are discussed and contrasted in the case study comparisons. For both lesson sequences, patterns observed as a result of activity sheet coding will be discussed and contrasted in the case study comparisons.

6. Delimitations of Data

a. Pre-Post Tests

Only the pre-post tests from matched sets of Whole class/Small group conditions (as described in this chapter in Sections B and C) are coded and analyzed statistically.

b. Videotapes and Transcripts

Although all videotapes that fit the criteria of the study were transcribed (33 videotapes) and all of those transcripts were read and subjected to some coding, only selected transcripts (4 per lesson topic, 8 total) were subjected to open coding. These were selected purposefully, according to whether they contained the phenomena for all coding categories to be developed. Many more of the transcripts were coded multiple times as part of the iterative honing process to create stable coding categories. Once stable countable codes were developed, these were applied to carefully delineated, but substantial, sections of all transcripts that were part of matched sets (19 transcripts in nine matched sets; some of these transcripts spanned multiple class periods). These transcript sections ranged from 12 to 62 minutes in length.

c. Student Activity Sheets

Only selected questions from the activity sheets were identified as useful for this study. Those sections were identified by examining a stratified sample of student work (30 activity sheets per topic) to see which questions appeared to have elicited sufficient student writing and drawing to have potential to address the research questions. Two or three questions per activity sheet were identified, as described in Chapter VI.

H. Addressing Questions of Validity and Reliability in Mixed Methods Research

In mixed methods research, as with all research, we want to make sure that we are actually evaluating what we think we are evaluating (validity) and that a given coding scheme will produce similar results each time it is applied to the data (reliability). These concepts have fairly accepted meanings when applied to quantitative analysis but their interpretation has been a matter of debate when applied to qualitative or mixed methods research.

A number of alternatives have been offered to the use of traditional interpretations of validity (Clement, 2000; Glaser 1978, 1998; Glaser & Strauss 1967; Merriam, 1998; Onwuegbuzie and Johnson, 2006). Rather than entering into that debate here, I have followed recommendations of Creswell (1998, 2003) and Clement (2000), and used multiple sources of data where possible to confirm or corroborate my findings. I have also used multiple methods of analysis to address the same questions. For instance, when looking for evidence that students had given or received visual support, I developed separate codes for visual and verbal evidence in the videotapes. To see whether students were actually using the visual features that were being supported, I looked for videotape evidence as well as evidence in their written work and then considered the results in light of each other. To help insure against researcher bias in the definition of these codes, I consulted with another knowledgeable professional. We particularly discussed borderline transcript episodes that could test the sensitivity and plausibility of the definitions. For the construction of the pre-post tests, teachers and other team members worked together to craft the questions to make sure we were addressing the concepts that we wished to address. To help insure that student responses were being interpreted as the students intended, the student answers to short answer questions, their written explanations to the explanation questions, and their drawn answers to the explanation questions were each analyzed separately. If this analysis revealed that students had interpreted a given question inconsistently or differently than intended, and that this had occurred often enough potentially to affect the statistical comparisons (generally, if it happened three or more times in a given class), the question was judged to have been ambiguous and was dropped from the study altogether. Another way of thinking about validity in the context of mixed methods research is to ask whether results are trustworthy. This was addressed by comparing the results of classes that had previously been matched along several parameters, as discussed above, and conducting the comparison blind to whether the class had been conducted in the whole class or small group condition.

Authors such as Clement (2000) and Denzin and Lincoln (2005) have suggested a rethinking of "reliability" in the context of exploratory research. Denzin and Lincoln have suggested the term "dependability" while Clement has discussed different levels of "observational reliability." A major part of the exploratory effort in this study went into generating new observational concepts for use in coding, a long and difficult process within the 'noise' and confusion of the classroom; this progressed slowly from initial, unstable coding concepts to stable concepts that could be applied consistently. Even in highly exploratory studies where training independent coders is not practical or appropriate (because the concepts used in the study take so long to develop), it is desirable that an evaluation or code, applied repeatedly, produces the same results, and that the coding criteria remain stable when applied to different data. To address this, I used several strategies. After coding categories were stable, I recoded earlier transcripts to see whether my own coding had been consistent over time, and if not, further refined the coding categories. After I applied the codes in a second context (a different lesson sequence), I revisited the earlier lesson sequence to see whether the codes had evolved or remained stable. This iterative process of code development continued until coding the same passage at different times produced the same results. At that point, another knowledgeable researcher coded a portion of the transcript passages or sample of the student work (about 10 student answers) to make sure there was substantial agreement. Any disagreement led to further refinement of the coding categories.

The hope is that enough information is given that a knowledgeable reader can judge for herself whether the conclusions appear plausible given the evidence, and that the methods are described well enough that they could be adapted for use in similar contexts by other interested researchers.

I. Summary of the Methodology

The analysis in this mixed methods study is conducted in three phases. Phase 1 involves the analysis of pre-post test gains with predominantly quantitative methods (Chapter V). This phase is designed to address the question of whether students are able to exhibit gains in conceptual reasoning in conjunction with the use of the simulations and animations in two modes: small group and whole class. Phase 2 uses mixed methods to analyze selected student activity sheet responses Chapter VI and video transcript data Chapter VII. This phase is designed to address the questions of whether students engage in reasoning about key concepts while working with the simulations, whether students recognize and use key features and potential affordances of the simulations, and what methods can be used to support the use of simulations and deal with student perceptual and conceptual difficulties. Phase 3, interwoven throughout Chapter VII, uses comparative case study analyses of matched sets of classes to compare countable codes from Phase 2 and to attempt to offer explanations for the results of Phases 1 and 2. The results from all matched sets are brought together to address each research question in Chapter VIII.

CHAPTER V

PRE-POST ANALYSIS

A. Pre-Post Test Analysis: Introduction

Because the classes chosen to be experimental and control groups were existing classes to which students had already been assigned by the schools, the participants in this study cannot be considered truly randomized. Rather, classes were selected to be part of matched sets according to whether they fit certain criteria in terms of age and general level of preparation of the students, amount of time spent on the lesson sequences, and comparable pre- and post-test times. (See Chapter IV.) In the absence of randomization, this study uses an identical pre-post test design. Comparisons of scores from pre-tests administered immediately before the lesson sequence provide an estimate of control and experimental group similarity. Rather than comparing raw post-test scores, gains from pre-tests to identical post-tests are used for experimental comparisons. Normalized gains (Hake, 1998) are also computed for the short answer questions because these gains involve another method of taking into account the variation in pre-test scores.

Gains were computed for all short answer questions and several selected explanation questions, as follows:

a) Raw Gain = $\dfrac{\text{(Post score – Pre score)}}{\text{(Maximum score)}}$

The scores on the explanation questions were not subjected to further analysis but are presented in tabular form. The scores on the short answer questions were subjected to further analysis as follows.

First, normalized (Hake) gains, which consider the amount of room for improvement between the pre-test results and a perfect score, were computed:

b) Hake gain = $\dfrac{\text{(Mean post score – Mean pre score)}}{\text{(Maximum score – Mean pre score)}}$

Independent samples *t*-tests were used to compare the pretest scores of the whole class and small group conditions within each matched set to see whether there had been a significant difference in pre-instruction performance between students in the two conditions. Paired samples *t*-tests were used to test the significance of gains within each matched set. Independent samples *t*-tests were used to compare gains across conditions within each matched set. For one matched set of three classes, an ANOVA was used to follow up on the results of the *t*-test. For all comparisons, O'Brien, Brown-Forsythe, Levene, and Barlett tests were used to assess the equality of variances of the samples. Effect sizes for all comparisons[7] were computed using Cohen's *d*, the difference in means divided by the pooled standard deviation (Cohen, 1992; Rosnow and Rosenthal, 2003).

$$\text{Cohen's } d = \frac{X_1 - X_2}{S_{\text{pooled}}}$$

$$S^2 = \frac{(n_1 - 1)*S_1^{\,2} + (n_2 - 1)*S_2^{\,2}}{n_1 + n_2 - 2}$$

where: d = Cohen's *d* effect size
X_1 = Mean gain of the whole class condition
X_2 = Mean gain of the small group condition
S = Pooled standard deviation
S_1 = Standard deviation for the whole class condition
S_2 = Standard deviation for the small group condition
n_1 = Number of subjects in the whole class condition
n_2 = Number of subjects in the small group condition

[7] Some authors recommend against reporting effect sizes in the absence of statistical difference. However, because the sample sizes are fairly small and I am concerned about the risk of Type II errors, I chose to follow the advice of authors such as Thompson (1999) and report effect sizes for all comparisons. These results and the *t*-test results will be considered in light of each other.

Cohen has suggested that effect sizes between 0.20 – 0.50 are small (which indicates that the difference in the means of the two samples is between 0,20 and 0,50 of a standard deviation), 0.50 - 0.80 are medium, and 0.80 and above are large. Below 0.20 is considered negligible.

B. Pre-Post Test Analysis: Gravitational Potential Energy

1. Short Answer Questions: Results

Scores were tabulated from short answer questions on the pre-post tests. (See example in Appendix C.) Gains are expressed as percentages of a perfect score. Abbreviations CP, HP, and AP refer to College Preparatory (lower level), Honors (mid level), and Advanced Placement (higher level) physics courses, respectively.

a. Gravitational Potential Energy Year 1 Honors Physics Teacher A

Table 5: Summary of Pre-Post Short Answer Analysis: Year 1 HP Teacher A

	Whole Class Condition	Small Group Condition
Mean Pre-Test Score	3.28 / 9.00 = 36.4%	4.25 / 9.00 = 47.2%
Mean Post-Test Score	5.28 / 9.00 = 58.7%	5.07 / 9.00 = 56.3%
Mean Pre-post Gain (Raw)	2.00 / 9.00 = 22.2%	0.82 / 9.00 = 9.1%
Mean Pre-post Gain (Hake)	2.00 / 5.72 = 35.0%	0.82 / 4.75 = 17.3%

In Teacher A's mid-level Honors classes, an independent samples *t*-test examining the difference in pretest scores of students in the two conditions resulted in a low *p*-value and a medium effect size, suggesting a difference between the two groups of students with respect to prior knowledge of the topics of the lesson [$t(37) = -1.65$, $p = 0.11$, $d = 0.53$]. Paired samples *t*-tests indicated that the gains were significant in both conditions, with a relatively large effect size for the whole class condition [$t(19) = 4.76$, $p < 0.001$*, $d = 1.13$] and a medium effect size for the small group condition [$t(18) = 2.62$, $p = 0.02$, $d = 0.51$]. An independent samples *t*-test comparing the pre-post gains in the two conditions revealed a significant difference due to condition with a medium effect size [$t(37) = 2.22$, $p = 0.03$*, $d = 0.71$]. Thus, the group that had a lower average pre-score had greater gains, and these resulted in similar post-test scores for the two groups. However, this did not appear to be the result of a ceiling effect. Figure 3 illustrates the mean scores and confidence intervals along with the pre-post gains.

Computing Hake gains revealed that the whole class condition achieved 35% of the gains possible for them, while the small group condition achieved only 17% of the gains possible for them. Although the N is small, there is no evidence here for a pre-post advantage for the small group condition. More details are included in Table 6.

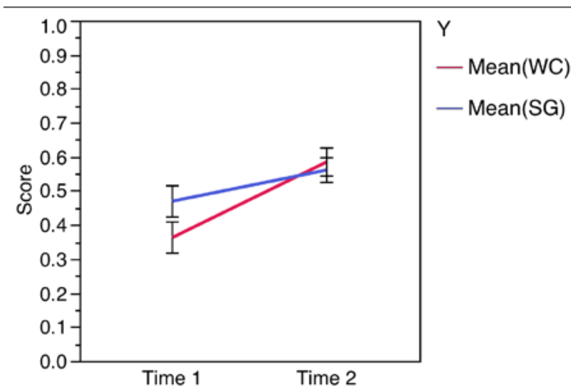
Table 6: Details of Pre-Post Short Answer Analysis: Year 1 HP Teacher A

Results of the paired samples t-tests, which examined changes in students' scores, pretest to posttest

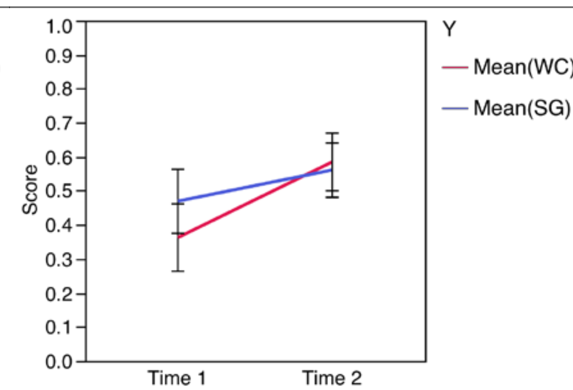
	Pre Mean (*SD*)	Post Mean (*SD*)	*t*-Value	df	Sig.	Cohen's *d*
WC (*N*=20)	0.36 (0.21)	0.59 (0.18)	4.759	19	< 0.001*	1.13
SG (*N*=19)	0.47 (0.19)	0.56 (0.16)	2.621	18	0.017*	0.51

Result of the independent samples t-test, which examined the difference in gains between the WC and SG conditions

	WC Gain Mean (*SD*)	SG Gain Mean (*SD*)	*t-Value*	df	Sig.	Cohen's *d*
	0.22 (0.21)	0.09 (0.15)	2.221	37	0.033*	0.71



a. Error bars one standard error from the mean.



b. Error bars show 95% confidence intervals.

Figure 3: Pre-Post Gains by Condition for Gravitational PE: Year 1 HP Teacher A

b. Gravitational Potential Energy Year 1 College Preparatory Physics Teacher B

Table 7: Summary of Pre-Post Short Answer Analysis: Year 1 CP Teacher B

	Whole Class Condition	Small Group Condition
Mean Pre-Test Score	2.58 / 9.00 = 28.7%	1.64 / 9.00 = 18.2%
Mean Post-Test Score	4.94 / 9.00 = 54.9%	3.93 / 9.00 = 43.7%
Mean Pre-post Gain (Raw)	2.36 / 9.00 = 26.2%	2.29 / 9.00 = 25.4%
Mean Pre-post Gain (Hake)	2.36 / 6.42 = 36.8%	2.29 / 7.36 = 31.1%

In Teacher B's lower-level College Preparatory classes, an independent samples *t*-test examining the difference in pretest scores of students in the two conditions resulted in a fairly low *p*-value and an effect size that was not negligible, suggesting that there may have been a difference between the two groups of students with respect to prior knowledge of the topics of the lesson [$t(23) = 1.19$, $p = 0.25$, $d = 0.48$]. Paired samples *t*-tests indicated that the gains were significant with relatively large effect sizes for the whole class condition [$t(10) = 4.45$, $p = 0.001$*, $d = 1.14$] and for the small group condition [$t(13) = 3.96$, $p = 0.002$*, $d = 1.40$]. Despite a possible prior difference hinted at by the pretest scores, an independent samples *t*-test comparing the pre-post gains in the two conditions revealed no significant difference in gains due to condition and a negligible effect size, size [$t(23) = 0.10$, $p = 0.92$, $d = 0.04$]. Figure 4 illustrates the mean scores and confidence intervals along with the pre-post gains.

Computing Hake gains revealed that the whole class condition achieved 37% of the gains possible for them, while the small group condition achieved 31% of the gains possible for them. As in the previous comparison, there is no evidence here for a pre-post advantage for the small group condition. See Table 8 for more details.

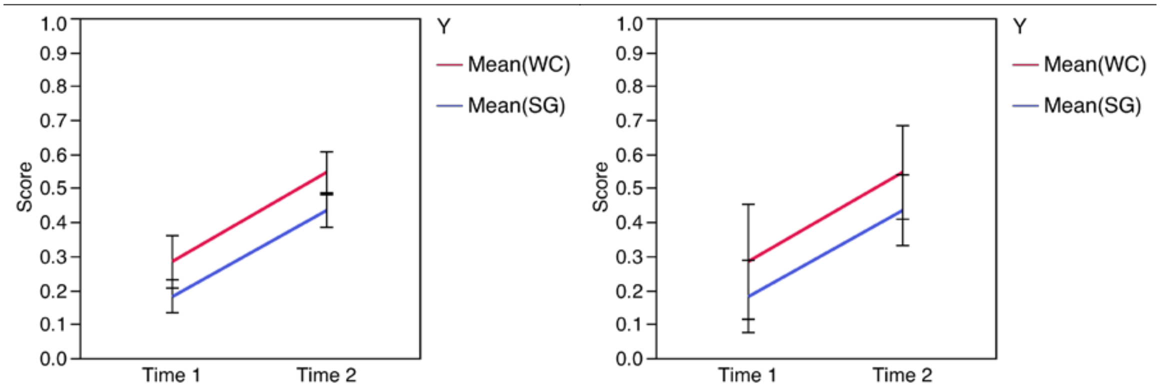
Table 8: Details of Pre-Post Short Answer Analysis: Year 1 CP Teacher B

Results of the paired samples t-tests, which examined changes in students' scores, pretest to posttest

	Pre Mean (*SD*)	Post Mean (*SD*)	*t*-Value	df	Sig.	Cohen's *d*
WC (*N*=11)	0.29 (0.25)	0.55 (0.21)	4.453	10	0.001*	1.14
SG (*N*=14)	0.18 (0.18)	0.44 (0.18)	3.960	13	0.002*	1.40

Result of the independent samples t-test, which examined the difference in gains between the WC and SG conditions

	WC Gain Mean (*SD*)	SG Gain Mean (*SD*)	*t-Value*	df	Sig.	Cohen's *d*
	0.26 (0.20)	0.25 (0.24)	0.097	23	0.924	0.04



a. Error bars one standard error from the mean. b. Error bars show 95% confidence intervals.

Figure 4: Pre-Post Gains by Condition for Gravitational PE: Year 1 CP Teacher B

c. Gravitational Potential Energy Year 1 Advanced Placement Physics Teacher B

Table 9: Summary of Pre-Post Short Answer Analysis: Year 1 AP Teacher B

	Whole Class Condition	Small Group Condition
Mean Pre-Test Score	6.83 / 9 = 75.9%	6.84 / 9 = 76.0%
Mean Post-Test Score	7.74 / 9 = 86.0%	7.02 / 9 = 78.0%
Mean Pre-post Gain (Raw)	0.91 / 9 = 10.1%	0.18 / 9 = 2.0%
Mean Pre-post Gain (Hake)	0.91 / 2.17 = 41.9%	0.18 / 2.16 = 8.3%

In Teacher B's higher-level Advanced Placement classes, the pretest means of the two conditions were almost identical; an independent samples *t*-test yielded no significant difference due to condition and a negligible effect size [$t(42) = -0.04$, $p = 0.97$, $d = 0.01$]. This suggested that the two groups of students were very similar with respect to prior knowledge of the topics of the lesson. However, paired samples *t*-tests revealed that the whole class condition had significant gains [$t(22) = 4.19$, $p < 0.001$*, $d = 0.60$] while the small group condition did not have gains significant at the $\alpha = 0.05$ level [$t(20) = 0.79$, $p = 0.44$, $d = 0.18$]. An independent samples *t*-test comparing the pre-post gains in the two conditions revealed a significant difference in favor of the whole class format with a medium effect size [$t(42) = 2.37$, $p = 0.02$*, $d = 0.71$]. Figure 5 illustrates the mean scores and confidence intervals along with the pre-post gains.

Although the gains were small, the difference in gains did not appear to be due to a ceiling effect. Computing Hake gains revealed that the whole class condition achieved 42% of the gains possible for them while the small group condition achieved only 8% of the gains possible for them. As in the previous two comparisons, there is no evidence here for a pre-post advantage for the small group condition. See Table 10 for more details.

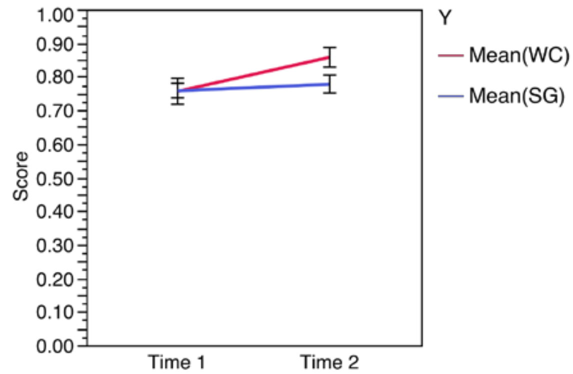
Table 10: Details of Pre-Post Short Answer Analysis: Year 1 AP Teacher B

Results of the paired samples t-tests, which examined changes in students' scores, pretest to posttest

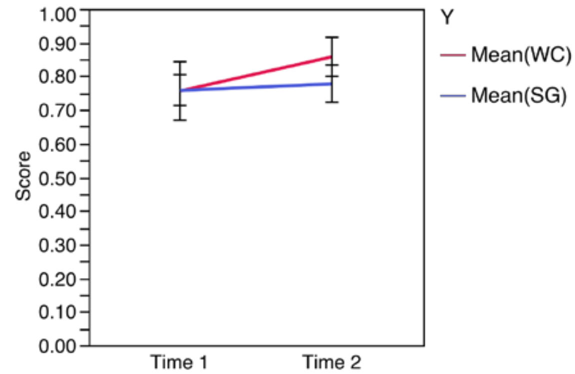
	Pre Mean (*SD*)	Post Mean (*SD*)	*t*-Value	df	Sig.	Cohen's *d*
WC (*N*=23)	0.76 (0.20)	0.86 (0.14)	4.188	22	< 0.001*	0.60
SG (*N*=21)	0.76 (0.10)	0.78 (0.12)	0.785	20	0.442	0.18

Result of the independent samples t-test, which examined the difference in gains between the WC and SG conditions

	WC Gain Mean (*SD*)	SG Gain Mean (*SD*)	*t-Value*	df	Sig.	Cohen's *d*
	0.10 (0.12)	0.02 (0.11)	2.368	42	0.023*	0.71



a. Error bars one standard error from the mean.



b. Error bars show 95% confidence intervals.

Figure 5: Pre-Post Gains by Condition for Gravitational PE: Year 1 AP Teacher B

d. Gravitational Potential Energy Year 2 Advanced Placement Physics Teacher B

Table 11: Summary of Pre-Post Short Answer Analysis: Year 2 AP Teacher B

	Whole Class Condition	Small Group Condition
Mean Pre-Test Score	5.74 / 9 = 63.8%	6.26 / 9 = 70.0%
Mean Post-Test Score	6.52 / 9 = 72.4%	6.86 / 9 = 76.2%
Mean Pre-post Gain (Raw)	0.78 / 9 = 8.7%	0.60 / 9 = 6.7%
Mean Pre-post Gain (Hake)	0.78 / 3.26 = 23.9%	0.60 / 2.74 = 21.9%

In the second year in Teacher B's higher-level Advanced Placement classes, an independent samples *t*-test examining the difference in pretest scores of students in the two conditions resulted in a fairly low *p*-value and an effect size that was not negligible [$t(40) = -1.36$, $p = 0.18$, $d = 0.42$], suggesting a possible difference between the two groups of students with respect to prior knowledge of the topics of the lesson. Paired samples *t*-tests indicated that the gains were significant in both conditions, with a medium effect size for the whole class condition [$t(20) = 2.43$, $p = 0.02$*, $d = 0.66$] and a small effect size for the small group condition [$t(20) = 3.07$, $p < 0.01$*, $d = 0.45$]. However, an independent samples *t*-test comparing the pre-post gains in the two conditions revealed no significant difference due to condition and a negligible effect size [$t(40) = 0.51$, $p =$

0.62, $d = 0.16$]. Figure 6 illustrates the mean scores and confidence intervals along with the pre-post gains.

Although the gains were small, there did not appear to be a ceiling effect on the post-test. In fact, computing Hake gains revealed that the whole class and small group conditions achieved only about a quarter of the gains possible for them, 24% and 22% respectively. As in the previous three comparisons, there is no evidence here for a pre-post advantage for the small group condition. See Table 12 for more details.

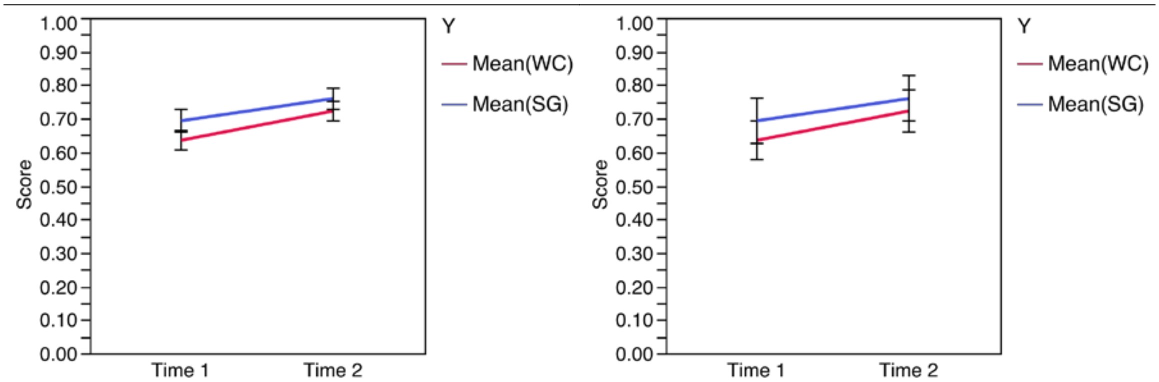
Table 12: Details of Pre-Post Short Answer Analysis: Year 2 AP Teacher B

Results of the paired samples t-tests, which examined changes in students' scores, pretest to posttest

	Pre Mean (*SD*)	Post Mean (*SD*)	*t*-Value	df	Sig.	Cohen's *d*
WC (*N*=21)	0.64 (0.12)	0.72 (0.14)	2.433	20	0.025*	0.66
SG (*N*=21)	0.70 (0.15)	0.76 (0.15)	3.068	20	0.006*	0.45

Result of the independent samples t-test assuming unequal variances[8]*, which examined the difference in gains between the WC and SG conditions*

	WC Gain Mean (*SD*)	SG Gain Mean (*SD*)	*t-Value*	df	Sig.	Cohen's *d*
	0.09 (0.16)	0.07 (0.10)	0.506	33	0.616	0.16



a. Error bars one standard error from the mean. b. Error bars show 95% confidence intervals.

Figure 6: Pre-Post Gains by Condition for Gravitational PE: Year 2 AP Teacher B

[8] Although O'Brien, Levene, and Barlett tests indicated the presence of a difference between the variances, assuming unequal variances did not change the *t*-statistic or *p*-value to within three significant digits. The only change was to the degrees of freedom.

2. Selected Explanation Questions: Coding Criteria and Results

The pre-post test was refined slightly before the second year and the explanation questions were reduced. All student answers were coded; this was done blind to condition (whole class or small group) and to time (pre or post). It became clear that some questions had appeared ambiguous to the students and these questions were dropped from the study. Two explanation questions on the Year 1 test and one on the Year 2 test remained. These asked whether a marble started from rest would reach the end of a track with a hill or a loop in it. The correct answers involved the amount of gravitational potential energy available at the start versus the kinetic energy required to get the marble past the obstacle in the middle. The loop problem, which appeared on both versions of the test, also required awareness of the centripetal acceleration needed at the top of the loop to keep the marble from falling. (See Appendix C.)]

These two explanation questions were coded either 0, ½, or 1 according to whether they were incorrect, partially correct, or correct. (See rubric at the end of Appendix C.) The results are not subjected to statistical analysis; rather, the average scores for the selected explanation questions are listed for each set of classes.

Table 13: Explanation Questions: Year 1 HP Teacher A
(Two explanation questions)
Concepts: For a bump on a track, considering the relative heights of bump and starting point is sufficient to determine whether a marble will make it over. For a loop in the track, centripetal acceleration must also be considered.

Condition	Mean 'Why' Pre	Mean 'Why' Post	Mean 'Why' Gain
WC (*N*=20)	29%	43%	**14%**
SG (*N*=19)	20%	46%	**26%**

Even though the whole class condition had performed significantly better on the short answer questions, the small group appeared to do better on the explanation questions. These are among the results that suggest the importance of case study analysis, which, among other things, can attempt to shed light on these scores.

Table 14: Explanation Questions: Year 1 CP Teacher B
(Two explanation questions)
Concepts: For a bump on a track, considering the relative heights of bump and starting point is sufficient to determine whether a marble will make it over. For a loop in the track, centripetal acceleration must also be considered.

Condition	Mean 'Why' Pre	Mean 'Why' Post	Mean 'Why' Gain
WC (*N*=11)	23%	27%	**5%**
SG (*N*=14)	20%	41%	**21%**

In this matched set, the whole class and small group conditions had almost identical gains on the short answer questions ($p < 0.92$), but the small group appeared to perform better on the explanation questions.

Table 15: Explanation Questions: Year 1 AP Teacher B
(Two explanation questions)
Concepts: For a bump on a track, considering the relative heights of bump and starting point is sufficient to determine whether a marble will make it over. For a loop in the track, centripetal acceleration must also be considered.

Condition	Mean 'Why' Pre	Mean 'Why' Post	Mean 'Why' Gain
WC (*N*=23)	35%	59%	**24%**
SG (*N*=21)	38%	40%	**2%**

In this matched set, the whole class condition did significantly better on the short answer questions, and also appears to have performed better on the explanation questions.

Table 16: Explanation Questions: Year 2 AP Teacher B
(One explanation question)
Concepts: For a loop in the track, considering the relative heights of loop and starting point is necessary but not sufficient to determine whether a marble will make it over; centripetal acceleration must also be considered.

Condition	Mean 'Why' Pre	Mean 'Why' Post	Mean 'Why' Gain
WC (*N*=21)	17%	60%	**43%**
SG (*N*=21)	7%	33%	**26%**

In the second year AP classes, the two conditions had similar gains on short answer questions, though these gains were small (less than 10%). However, the whole class condition appears to have performed better on the explanation question. These results suggest questions for case study analysis concerning the nature of the changes the teacher made for the Year 2 lesson in response to her Year 1 experiences.

C. Pre-Post Test Analysis: Projectile Motion

1. Short Answer Questions: Results

Scores were tabulated from short answer questions on the pre-post tests as described in the introduction to this chapter. All gains are expressed as percentages of a perfect score. Abbreviations CP, HP, and AP refer to College Preparatory (lower level), Honors (mid level), and Advanced Placement (higher level) Physics courses, respectively.

a. Projectile Motion Year 1 Honors Physics Teacher A

Table 17: Summary of Pre-Post Short Answer Analysis: Year 1 HP Teacher A

	WC Condition	SG Condition
Mean Pre-Test Score	4.29 / 8.00 = 53.6%	3.08 / 8.00 = 38.5%
Mean Post-Test Score	7.10 / 8.00 = 88.8%	5.96 / 8.00 = 74.5%
Mean Pre-post Gain (Raw)	2.81 / 8.00 = 35.1%	2.88 / 8.00 = 36.0%
Mean Pre-post Gain (Hake)	2.81 / 3.71 = 75.7%	2.88 / 4.92 = 58.5%

In Teacher A's mid-level Honors classes, an independent samples *t*-test examining the difference in pretest scores of students in the two conditions resulted in a low *p*-value and a medium effect size [$t(44) = 1.76$, $p = 0.09$, $d = 0.52$], suggesting some difference between the two groups of students with respect to prior knowledge of the topics of the lesson. Paired samples *t*-tests indicated that the gains were significant with relatively large effect sizes for both the whole class condition [$t(21) = 4.65$, $p < 0.001$*, $d = 1.33$] and the small group condition [$t(25) = 5.34$, $p < 0.001$*, $d = 1.17$]. Despite the apparent difference in pretest scores, an independent samples *t*-test comparing the pre-post gains in the two conditions revealed no significant difference due to condition and a negligible effect size [$t(44) = -0.09$, $p = 0.93$, $d = 0.03$]. Figure 7 illustrates the mean scores and confidence intervals along with the pre-post gains.

Computing Hake gains revealed that the whole class condition achieved 76% of the gains possible for them, while the small group condition achieved only 59% of the gains possible for them. In other words, even though the two groups had almost identical gains, the whole class group achieved almost all of the gains possible for them. Although there could be many reasons for this, it seems clear that there is no evidence here for a pre-post advantage for the small group condition. See Table 18 for more details.

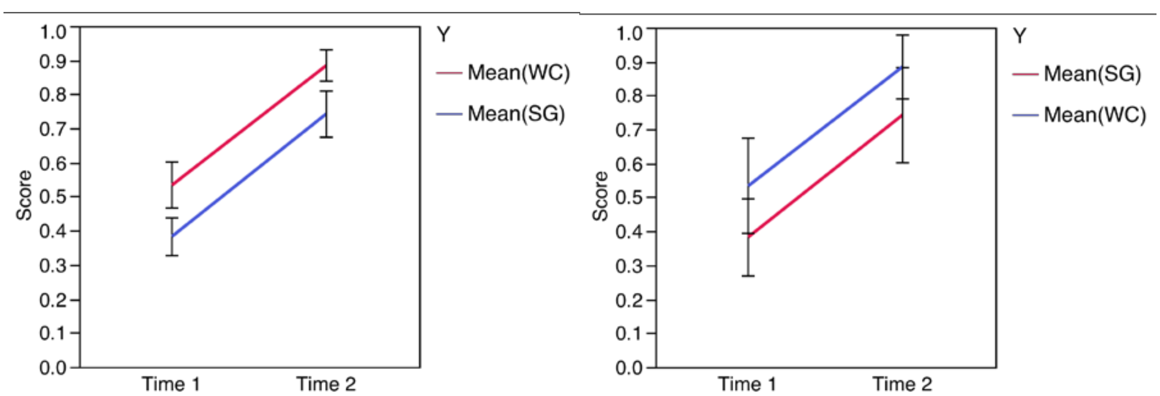
Table 18: Details of Pre-Post Short Answer Analysis: Year 1 HP Teacher A

Results of the paired samples t-tests, which examined changes in students' scores, pretest to posttest

	Pre Mean (*SD*)	Post Mean (*SD*)	*t*-Value	df	Sig.	Cohen's *d*
WC (*N*=21)	0.54 (0.31)	0.89 (0.21)	4.651	20	< 0.001*	1.33
SG (*N*=25)	0.39 (0.27)	0.75 (0.34)	5.338	24	< 0.001*	1.17

Result of the independent samples t-test, which examined the difference in gains between the WC and SG conditions

	WC Gain Mean (*SD*)	SG Gain Mean (*SD*)	*t-Value*	df	Sig.	Cohen's *d*
	0.35 (0.35)	0.36 (0.34)	-0.087	44	0.931	0.03



a. Error bars one standard error from the mean. b. Error bars show 95% confidence intervals.

Figure 7: Pre-Post Gains by Condition for Projectile Motion: Year 1 HP Teacher A

b. Projectile Motion Year 1 Honors Physics Teacher C

Table 19: Summary of Pre-Post Short Answer Analysis: Year 1 HP Teacher C

	WC Condition	SG Condition
Mean Pre-Test Score	4.15 / 8.00 = 51.9%	4.11 / 8.00 = 51.4%
Mean Post-Test Score	6.91 / 8.00 = 86.4%	6.68 / 8.00 = 83.5%
Mean Pre-post Gain (Raw)	2.76 / 8.00 = 34.6%	2.57 / 8.00 = 32.1%
Mean Pre-post Gain (Hake)	2.76 / 3.85 = 71.7%	2.57 / 3.89 = 66.1%

In Teacher C's mid-level Honors classes (two classes used the whole class format and one used the small group format), an independent samples *t*-test examining the pretest scores yielded no significant difference due to condition and a negligible effect size [$t(51) = 0.08$, $p = 0.94$, $d = 0.02$]. This suggested that the two groups of students were very similar with respect to prior knowledge of the topics of the lesson. Paired samples *t*-tests indicated that the gains were significant with relatively large effect sizes for both the whole class condition [$t(33) = 6.86$, $p < 0.001$*, $d = 1.43$] and the small group condition [$t(18) = 5.84$, $p < 0.001$*, $d = 1.63$]. An independent samples *t*-test comparing the pre-post gains in the two conditions revealed no significant difference due to condition and a negligible effect size [$t(51) = 0.29$, $p = 0.77$, $d = 0.08$]. Because of concern about the possibility of a Type II error (the failure to identify a difference where one exists), a one-way ANOVA was conducted among the three classes; it found no difference in gains among the classes [$F(2, 50) = 0.06$, $p = 0.94$]. Figure 8 illustrates the mean scores and confidence intervals along with the pre-post gains (both by condition and by class).

Computing Hake gains revealed that the whole class condition achieved 72% of the gains possible for them, and the small group condition achieved 66% of the gains possible for them. As with Teacher A's HP classes, there is again no evidence for a pre-post advantage for the small group condition. See Table 20 for more details.

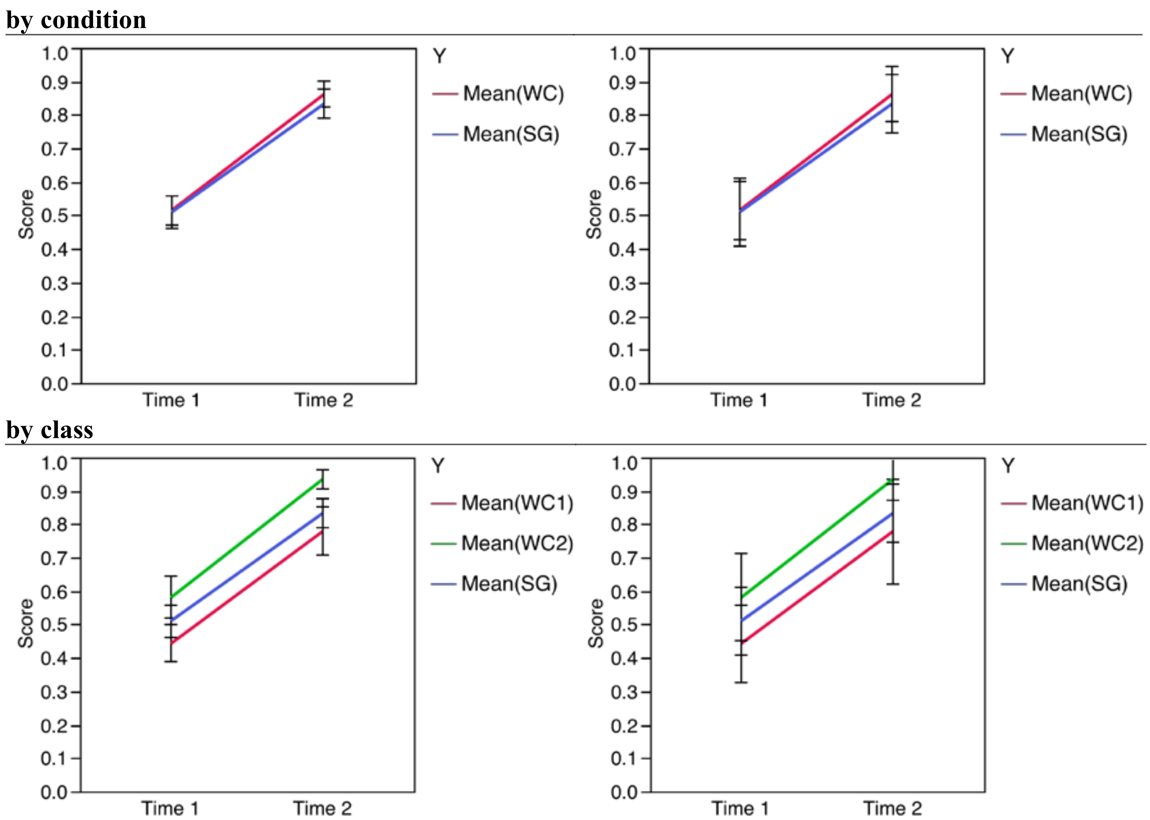
Table 20: Details of Pre-Post Short Answer Analysis: Year 1 HP Teacher C

Results of the paired samples t-tests, which examined changes in students' scores, pretest to posttest

	Pre Mean (*SD*)	Post Mean (*SD*)	*t*-Value	df	Sig.	Cohen's *d*
WC (*N*=34)	0.52 (0.25)	0.86 (0.23)	6.862	33	< 0.001*	1.43
SG (*N*=19)	0.51 (0.21)	0.84 (0.18)	5.843	18	< 0.001*	1.63

Result of the independent samples t-test, which examined the difference in gains in the WC and SG conditions

	WC Gain Mean (*SD*)	SG Gain Mean (*SD*)	*t-Value*	df	Sig.	Cohen's *d*
	0.35 (0.29)	0.32 (0.24)	0.294	51	0.770	0.08



a. Error bars one standard error from the mean. b. Error bars show 95% confidence intervals.

Figure 8: Pre-Post Gains by Condition and Class for Projectile Motion: Year 1 HP Teacher C

To protect against the possibility of a Type II error, a comparison was run among the three classes as well as between conditions.

c. Projectile Motion Year 2 Honors Physics Teacher A

Table 21: Summary of Pre-Post Short Answer Analysis: Year 2 HP Teacher A

	Whole Class Condition	Small Group Condition
Mean Pre-Test Score	2.83 / 8 = 35.4%	2.66 / 8 = 33.2%
Mean Post-Test Score	6.13 / 8 = 76.7%	5.59 / 8 = 69.9%
Mean Pre-post Gain (Raw)	3.30 / 8 = 41.3%	2.93 / 8 = 36.6%
Mean Pre-post Gain (Hake)	3.30 / 5.17 = 63.8%	2.93 / 5.34 = 54.9%

In Teacher A's Year 2 Honors level classes, an independent samples *t*-test examining the pretest scores yielded no significant difference due to condition and a negligible effect size [$t(35) = 0.24$, $p = 0.81$, $d = 0.08$]. This suggested that the two groups of students were similar with respect to prior knowledge of the topics of the lesson. Paired samples *t*-tests indicated that the gains were significant with relatively large effect sizes for both the whole class condition [$t(14) = 7.22$, $p < 0.001$*, $d = 1.37$] and the small group condition [$t(21) = 5.20$, $p < 0.001$*, $d = 1.23$]. An independent samples *t*-test comparing the pre-post gains in the two conditions revealed no significant difference due to condition and a negligible effect size [$t(35) = 0.47$, $p = 0.64$, $d = 0.16$]. Figure 9 illustrates the mean scores and confidence intervals along with the pre-post gains.

Computing Hake gains revealed that the whole class condition achieved 64% of the gains possible for them and the small group condition 55% of the gains possible for them. This difference, while not huge, is of some interest because the mean pre scores were so similar. As with the first two Projectile Motion comparisons, there is no evidence for a pre-post advantage for the small group condition. See Table 22 for more details.

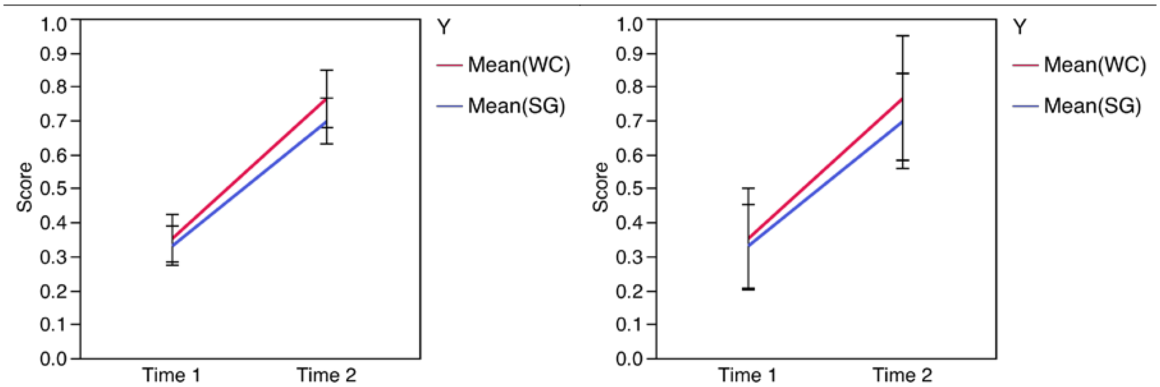
Table 22: Details of Pre-Post Short Answer Analysis: Year 2 HP Teacher A

Results of the paired samples t-tests, which examined changes in students' scores, pretest to posttest

	Pre Mean (*SD*)	Post Mean (*SD*)	*t*-Value	df	Sig.	Cohen's *d*
WC (*N*=15)	0.35 (0.27)	0.77 (0.33)	7.218	14	< 0.001*	1.37
SG (*N*=22)	0.33 (0.28)	0.70 (0.32)	5.204	21	< 0.001*	1.23

Result of the independent samples t-test, which examined the difference in gains in the WC and SG conditions

	WC Gain Mean (*SD*)	SG Gain Mean (*SD*)	*t-Value*	df	Sig.	Cohen's *d*
	0.41 (0.22)	0.37 (0.33)	0.471	35	0.640	0.16



a. Error bars one standard error from the mean. b. Error bars show 95% confidence intervals.

Figure 9: Pre-Post Gains by Condition for Projectile Motion: Year 2 HP Teacher A

<u>d. Projectile Motion Year 2 College Preparatory Physics Teacher A</u>

Table 23: Summary of Pre-Post Short Answer Analysis: Year 2 CP Teacher A

	Whole Class Condition	Small Group Condition
Mean Pre-Test Score	2.21 / 8 = 27.6%	2.67 / 8 = 33.4%
Mean Post-Test Score	4.71 / 8 = 58.9%	4.83 / 8 = 60.4%
Mean Pre-post Gain (Raw)	2.50 / 8 = 31.2%	2.16 / 8 = 27.0%
Mean Pre-post Gain (Hake)	2.50 / 5.79 = 43.2%	2.16 / 5.33 = 40.5%

In Teacher A's lower-level College Preparatory classes, an independent samples *t*-test examining the pretest scores of students yielded no significant difference due to condition and a small effect size [$t(21) = -0.50$, $p = 0.62$, $d = 0.22$]. This suggested that the two groups of students were similar with respect to prior knowledge of the topics of the lesson. Paired samples *t*-tests indicated that the gains were significant with relatively large effect sizes for both the whole class condition [$t(13) = 3.99$, $p = 0.002$*, $d = 1.29$] and the small group condition [$t(8) = 2.84$, $p = 0.022$*, $d = 0.89$]. An independent samples *t*-test comparing the pre-post gains in the two conditions revealed no significant difference due to condition and a negligible effect size [$t(21) = 0.34$, $p = 0.74$, $d = 0.14$]. Figure 10 illustrates the mean scores and confidence intervals along with the pre-post gains.

Computing Hake gains revealed that the whole class condition achieved 43% of the gains possible for them, and the small group condition 41% of the gains possible for them. As with the first three Projectile Motion comparisons, there is no evidence for a pre-post advantage for the small group condition. See Table 24 for more details.

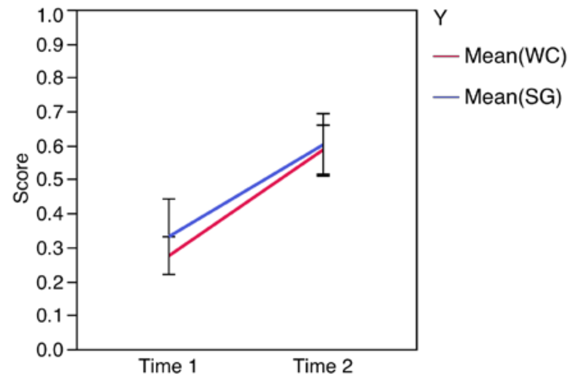
Table 24: Details of Pre-Post Short Answer Analysis: Year 2 CP Teacher A

Results of the paired samples t-tests, which examined changes in students' scores, pretest to posttest

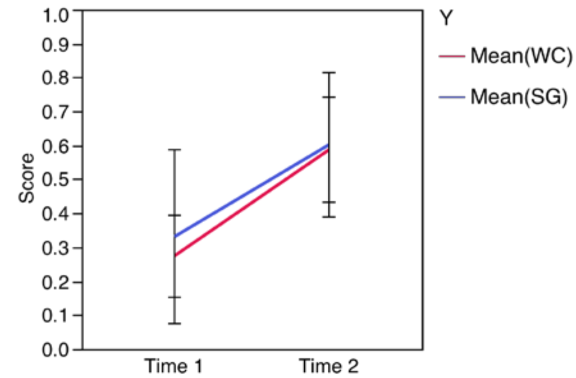
	Pre Mean (*SD*)	Post Mean (*SD*)	*t*-Value	df	Sig.	Cohen's *d*
WC (*N*=14)	0.28 (0.21)	0.59 (0.27)	3.989	13	0.002*	1.29
SG (*N*=9)	0.33 (0.33)	0.60 (0.28)	2.837	8	0.022*	0.89

Result of the independent samples t-test, which examined the difference in gains in the WC and SG conditions

	WC Gain Mean (*SD*)	SG Gain Mean (*SD*)	*t-Value*	df	Sig.	Cohen's *d*
	0.31 (0.29)	0.27 (0.29)	0.336	21	0.741	0.14



a. Error bars one standard error from the mean.



b. Error bars show 95% confidence intervals.

Figure 10: Pre-Post Gains by Condition for Projectile Motion: Year 2 CP Teacher A

e. Projectile Motion Year 2 Advanced Placement Physics Teacher B

Table 25: Summary of Pre-Post Short Answer Analysis: Year 2 AP Teacher B

	Whole Class Condition	Small Group Condition
Mean Pre-Test Score	4.00 / 6 = 66.7%	3.93 / 6 = 65.5%
Mean Post-Test Score	5.33 / 6 = 88.8%	5.24 / 6 = 87.3%
Mean Pre-post Gain (Raw)	1.33 / 6 = 22.2%	1.31 / 6 = 21.8%
Mean Pre-post Gain (Hake)	1.33 / 2.00 = 66.5%	1.31 / 2.07 = 63.3%

In Teacher B's higher-level Advanced Placement classes, an independent samples *t*-test examining the pretest scores yielded no significant difference due to condition and a negligible effect size [$t(39) = 0.13$, $p = 0.89$, $d = 0.04$]. This suggested that the two groups of students were very similar with respect to prior knowledge of the topics of the lesson. Paired samples *t*-tests indicated that the gains were significant with relatively large effect sizes for both the whole class condition [$t(19) = 4.42$, $p < 0.001$*, $d = 0.89$] and the small group condition [$t(20) = 4.34$, $p < 0.001$*, $d = 1.01$]. An independent samples *t*-test comparing the pre-post gains in the two conditions revealed no significant difference due to condition and a negligible effect size [$t(39) = 0.04$, $p = 0.97$, $d = 0.01$]. Figure 11 illustrates the mean scores and confidence intervals along with the pre-post gains.

Computing Hake gains revealed that the whole class condition achieved 67% of the gains possible for them and the small group condition 63% of the gains possible for them. As with the first three Projectile Motion comparisons, there is no evidence for a pre-post advantage for the small group condition. See Table 26 for more details.

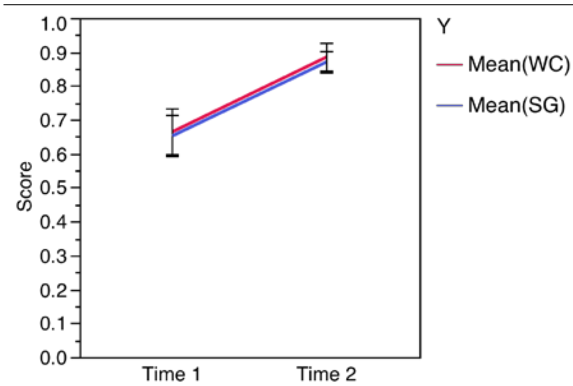
Table 26: Details of Pre-Post Short Answer Analysis: Year 2 AP Teacher B

Results of the paired samples t-tests, which examined changes in students' scores, pretest to posttest

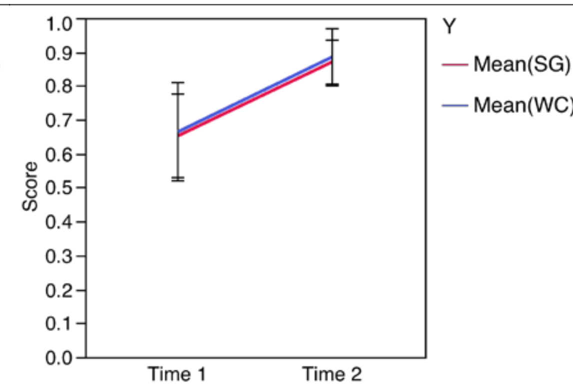
	Pre Mean (*SD*)	Post Mean (*SD*)	*t*-Value	df	Sig.	Cohen's *d*
WC (*N*=20)	0.67 (0.30)	0.89 (0.18)	4.421	19	< 0.001*	0.89
SG (*N*=21)	0.65 (0.27)	0.87 (0.15)	4.340	20	< 0.001*	1.01

Result of the independent samples t-test, which examined the difference in gains in the WC and SG conditions

	WC Gain Mean (*SD*)	SG Gain Mean (*SD*)	*t-Value*	df	Sig.	Cohen's *d*
	0.22 (0.22)	0.22 (0.23)	0.036	39	0.971	0.01



a. Error bars one standard error from the mean.



b. Error bars show 95% confidence intervals.

Figure 11: Pre-Post Gains by Condition for Projectile Motion: Year 2 AP Teacher B

2. Selected Explanation Questions: Coding Criteria and Results

Although the pre-post tests varied with physics level and year, each test had at least two questions that asked for further explanation. All student answers were coded; this was done blind to condition (whole class or small group) and to time (pre or post). It became clear that some questions had appeared ambiguous to the students and these questions were dropped from the study. Two explanation questions per test remained. These addressed aspects of projectile motion: change of vertical and horizontal components of velocity with change in gravity; trade off between hang time and ground speed in determining range; and effect on hang time and range due to a small variation in angle. (See Appendix G.) The explanation questions were coded either 0, ½, or 1 according to whether they were incorrect, partially correct, or correct. (See sample rubric at end of Appendix G.)

Table 27: Explanation Questions: Year 1 HP Teacher A
(Two explanation questions)
Concept: If the value of gravity changes, the vertical component of velocity will change but the horizontal component will not be affected.

Condition	Mean 'Why' Pre	Mean 'Why' Post	Mean 'Why' Gain
WC (*N*=21)	44%	63%	**19%**
SG (*N*=25)	33%	44%	**11%**

The whole class condition appears to have performed better than the small group condition on the explanation questions. As can be seen in the previous section, the gains of the whole class and small group conditions for the short answer questions were almost identical ($p < 0.93$, effect size $d = 0.03$). Case study analyses will seek to shed light on these scores.

Table 28: Explanation Questions: Year 1 HP Teacher C
(Two explanation questions)
Concept: If the value of gravity changes, the vertical component of velocity will change but the horizontal component will not be affected.

Condition	Mean 'Why' Pre	Mean 'Why' Post	Mean 'Why' Gain
WC 1 (*N*=16)	30%	39%	**9%**
WC 2 (*N*=18)	35%	58%	**24%**
SG (*N*=19)	30%	42%	**12%**

Among Teacher C's three classes, as discussed in the previous section, there was no statistical difference in gains on the short answer questions. However, the fact that the gains for the explanation questions appear to follow the same pattern as did the gains for the short answer questions (WC 2 had the highest gains and WC 1 the lowest), and the fact that the Hake gains followed this same pattern, suggests a possible trend among the classes—though, once again, there is no evidence here of an advantage for the small group format.

Table 29: Explanation Questions: Year 2 HP Teacher A
(Two explanation questions)
Concept: For a given launch angle and launch velocity, the range is determined by a trade-off between hang time and ground speed. Complementary angles produce the same range.

Condition	Mean 'Why' Pre	Mean 'Why' Post	Mean 'Why' Gain
WC (*N*=15)	8%	25%	**17%**
SG (*N*=22)	11%	17%	**6%**

In this matched set, there was no significant difference between whole class and small group conditions for gains on the short answer questions. However, the whole class condition appears to have performed better on the explanation questions.

Table 30: Explanation Questions: Year 2 CP Teacher A
(Two explanation questions)
Concept: For a given launch angle and launch velocity, the range is determined by a trade-off between hang time and ground speed. Complementary angles produce the same range.

Condition	Mean 'Why' Pre	Mean 'Why' Post	Mean 'Why' Gain
WC (*N*=14)	4%	14%	**11%**
SG (*N*=9)	8%	3%	**-6%**

For this lower-level class, there was no significant difference between whole class and small group conditions on the short answer questions, but there may have been a difference in performance on the explanation questions. Interestingly, the small group condition does not appear to have performed any better on the explanation questions on the post-test than on the pre-test.

Table 31: Explanation Questions: Year 2 AP Teacher B
(Two explanation questions)
Concept: For two projectiles launched at slightly different angles, the angle closer to 90 degrees will produce the longest hang time and the angle closer to 45 degrees will produce the longest range.

Condition	Mean 'Why' Pre	Mean 'Why' Post	Mean 'Why' Gain
WC (*N*=20)	38%	56%	**19%**
SG (*N*=21)	44%	61%	**17%**

These two higher-level classes had almost identical gains for the short answer questions ($p < 0.97$, $d = 0.01$) and appear to have performed very similarly on the explanation questions as well.

All of these results will be discussed further in Chapters VII and VIII.

CHAPTER VI

ACTIVITY SHEET ANALYSIS: SELECTED QUESTIONS

A. Activity Sheet Analysis: Introduction

The aim of the activity sheet analysis was to address Research Question 5, about student use of certain key features of the simulations and animations to be described below. As will be seen in Chapters VII and VIII, videotape analysis revealed that at least some students mentioned the key features during each discussion where this aspect was analyzed. However videotape analysis, by itself, proved not to be sufficient to address the question of whether students actually used the key features in their own thinking. Pre-post tests did not address this issue because they consisted of transfer questions (questions that related conceptually to the activity sheet questions but that did not use the same scenarios). I examined student work on the activity sheets to see whether evidence for use of the key features could be identified there. The activity sheets, as with the pre-post tests, gave an opportunity to look at the work of every student in a way that videotape analysis cannot do; on the other hand, some students wrote very briefly or drew their answers, making it a challenge to interpret their work.

B. Activity Sheet Analysis: Gravitational Potential Energy Lessons

1. Selected Gravitational Potential Energy Activity Sheet Explanation Questions: Coding Criteria

For Gravitational Potential Energy, the simulation used was Energy Skate Park from the PhET simulations (Reid, et al., 2009). The features that had been identified as key in exploratory studies (Chapter III) were the Gravitational Potential Energy (GPE) reference line and the Energy Bar Graph.

For this one or two-day lesson, the activity sheets varied slightly from year to year and teacher to teacher, but this was not of concern because the small group and whole class conditions to be compared within each matched set always used the same versions. I examined ten examples of student work selected at random from each version of the activity sheet, five from whole class and five from small group conditions. This was done blind to condition. In this sample of 30 activity sheets, almost all of the references to the key features were made in response to the questions below:

Activity sheet (versions 1 & 2)
Question 7: Could the total energy be zero at some position? Explain.

Activity sheet (version 3)
Question 6b: Could potential energy ever be less than zero?
Question 7: Could total energy ever be less than zero? Explain.

In Version 3, the teacher had modified Question 7, which originally asked about $TE = 0$, to ask instead about $TE < 0$. She also created a new Question 6 to lead up to Question 7, asking whether *KE* or *PE* could ever be less than zero. The part of Question 6 of interest in the present context was 6b, which asked about gravitational potential energy. Some students using Version 3 wrote their answer to Question 7 as a continuation of their answer to Question 6b, so it seemed reasonable and practical to consider the two answers as a single unit.

In consultation with a knowledgeable expert, I decided to use open coding on all activity sheets for student responses to Q6b (version 3) and Q7 (all versions) about whether energy in a system could equal a zero or negative amount. This coding was done blind to whole class or small group condition.

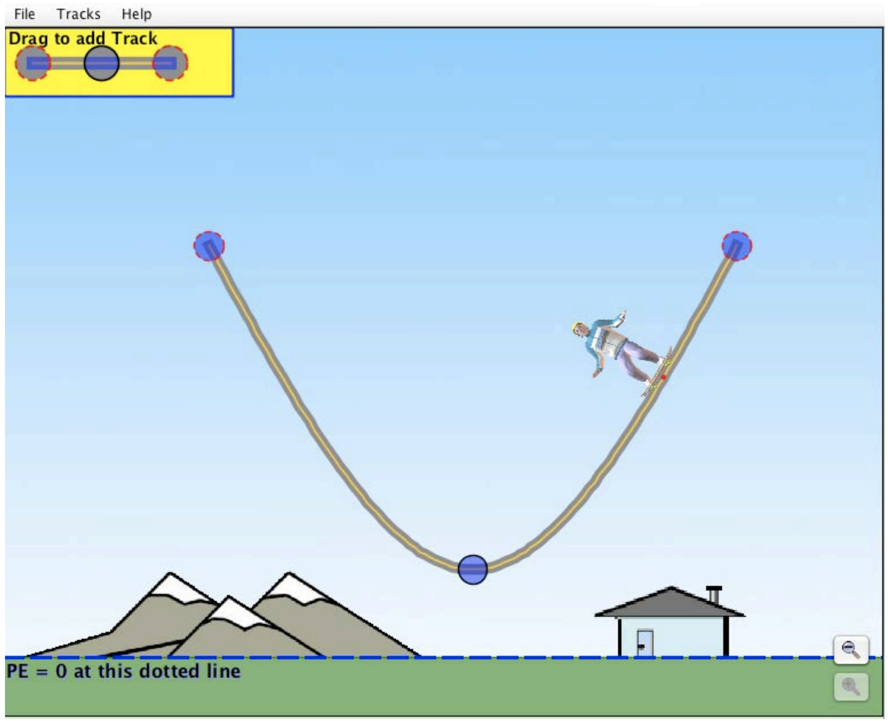


Figure 12: The Gravitational Potential Energy Reference line.
The GPE line at its default position in Energy Skate Park (PhET). Some students mistakenly identified the GPE line as synonymous with the surface of the ground, h=0.

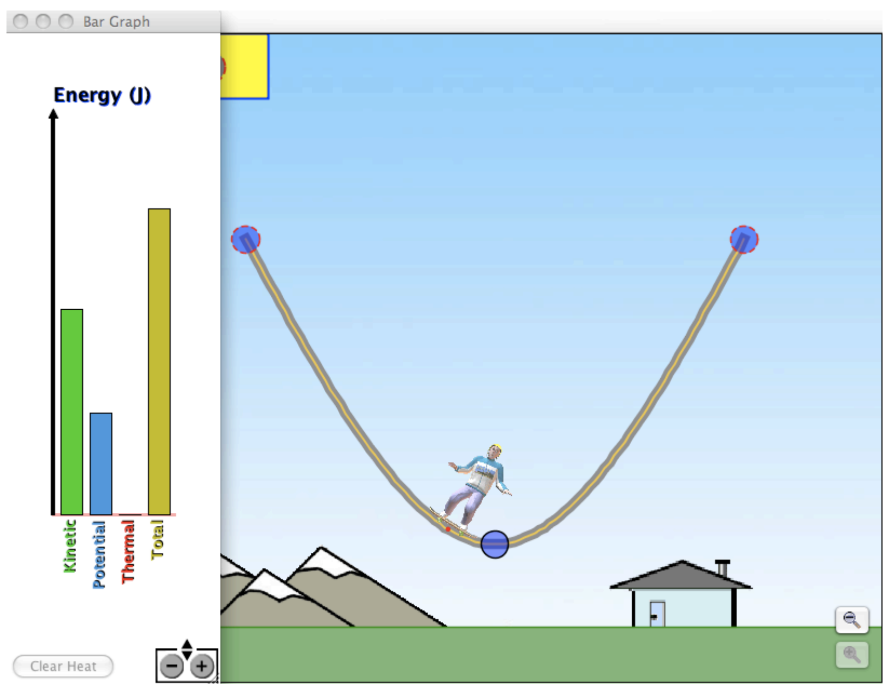


Figure 13: The Energy Bar Graph.
The bar graph appeared to support a focus on semi-quantitative thinking and recognition of dynamic relationships between changing energy values. There were two additional animated energy graphs that, unlike the bar graph, included numerical information on the axes; during classroom observation, these appeared to encourage computation.

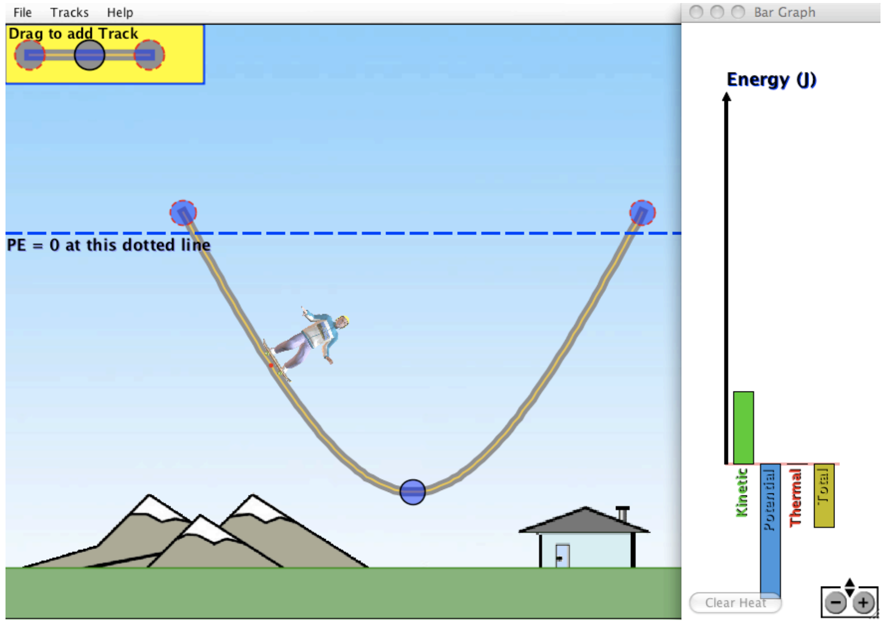


Figure 14: GPE line and Bar Graph used together.
In this screen shot, the bar graph shows *negative* energy amounts. This could occur in the simulation only if the GPE line had been moved away from its default position.

2. Evidence for Student Use of GPE Reference Line

In some videotapes, there were episodes during discussion when the reference line feature was turned on and was moved around. In other videotapes, it appeared that the reference line was never discovered or was discovered but never moved away from its default position on the ground. (Compare Figures 12 through 14 above.) In situations where it was never discovered, some students turned on an optional grid feature and referred instead to an immovable grid line at ground level labeled $h = 0$. This was problematic because an important point of the lesson was that the zero line for potential energy can be chosen arbitrarily.

In student written responses, it was often not possible to tell with certainty whether a student was referring to a reference line that was movable and denoted an arbitrary zero point, or whether they were referring to a fixed line along the ground at $h =$

0. Some students explicitly mentioned a "reference line" or "zero bar," while others mentioned "$h = 0$" or "the ground." I divided the responses into those that appeared to refer to something movable, such as "a bar," and those that appeared to refer to something immovable, such as "the ground" or "the bottom." I regard the first category as providing some evidence that the student did employ the idea of a movable reference line as they answered the question, and the second category as providing insufficient evidence for this. A third category, of answers that did not mention a line or zero point, was collapsed into the second category. This produced two mutually exclusive categories of answers, one where the answers provided some evidence that the student employed the idea of a movable reference line, and the other where the answers provided insufficient or no evidence. Sample phrases from student responses are listed in Table 32.

Table 32: Does student refer to the GPE reference line in a way that implies that the line is movable?

Yes = 1 **(sufficient evidence)**	**No = 0** **(insufficient or no evidence)**
"When the object is below the ref line"	"When h is always negative"
"In our setup with the line at the bottom, TE stays the same."	"When not moving at $h = 0$"
"If you put the zero bar at the top where he stops momentarily"	"If he's at the bottom, not moving"

3. Evidence for Student Use of Animated Energy Graph

I wished to categorize the student answers to Q6 (version 3) and Q7 (all versions) according to whether they contained evidence that the students had used the animated energy graph in their thinking. The development of the above categories for use of the GPE line raised the possibility that a similar framework could be developed concerning use of the energy bar graph. However, the open coding process revealed that not a single student response on the activity sheets explicitly mentioned the energy bar graph (or any other graph). This posed a challenge. As with the GPE reference line, videotapes had revealed considerable variation in how often students attended to the energy graph. Teachers varied in how often they referred to this feature, and in one whole class discussion, it was not clear that the teacher had referred to it at all. (It *was* referred to in all versions of the activity sheet and students were instructed to use it.)

Although the energy graphs were not mentioned, patterns did appear in the answers with respect to how energy quantities were referred. The below list of answers (composites of actual student answers to whether *TE* could be zero at some position or whether *KE*, *PE*, or *TE* could ever be less than zero) can be used to illustrate the patterns that were identified.

1. "*TE* can never equal 0 because energy is conserved."
2. "*TE* can never change."
3. "No because there is always some kind of energy."
4. "Only if he were on the ground not moving."
5. "If *KE* + *PE* = 0."
6. "If *KE* + *PE*= -*TE*"
7. "No, because if you have no *KE*, then you have *PE*, and vice versa."
8. "*TE* can go negative when he goes lower than the reference line."
9. "If *PE* is negative and larger than *KE*."

In a number of answers similar to #1 and #2, it appeared that students were applying the principle of conservation of energy beyond the domain of application of this principle. (This is consistent with episodes on the videotapes in which students were heard saying that the energy in the system would not change even if one moved the system from the Earth to the Moon because "*TE* can never change.") In answers similar to #3 and #4, it appeared that students believed that in order for *TE* to equal zero, each and every kind of energy would have to equal zero. Therefore, the skater would have to be still and positioned at $h = 0$. Some students mentioned that thermal energy would also have to equal zero. It can be hypothesized that these students did not have a concept of negative energy and so did not consider the possibility that *KE* and *PE* could cancel and add to zero. The fifth and sixth answers exhibit a formal understanding of the scientifically accepted answer. Although the seventh answer is incorrect, it exhibits an awareness of a relationship between *KE* and *PE* on the half-pipe. The eighth answer exhibits an awareness of a relationship between the reference line and the amount of *TE*. The ninth answer exhibits awareness that *PE* can take on negative values and that it can overbalance *KE* (which is always positive) in order to result in a negative *TE*. This is similar to the idea expressed in answer #6, although in answer #9 it is expressed in words rather than via a formula.

It is of interest that the last three answers are consistent with a description of the motion of the energy bars in the animated energy bar graph; these answers do not appear to derive from formalism. The first four answers are not consistent with such a description; they do not describe what students would have seen had they attended to the movements of the bars in the animated energy bar graph. Answers 5 and 6 are not inconsistent with the motion of the energy bars but are expressed purely as formalism; they could have come from rote memorization without understanding, similar to the way some students in these classes quoted the phrase "Conservation of Energy" without exhibiting understanding of what it meant.

The concepts suggested by the last three answers include: *PE* and *KE* change relative to each other, when one goes up, the other goes down; *TE* reading for a system can change (and this does not require changing a physical aspect of the system); *PE* and/or *TE* quantities can become negative.

These are precisely the concepts about changing energy that the energy bar exercises were designed to support, and the third concept was identified as a key concept for this lesson. Student written answers do not show for certain where the concepts were acquired. But what can be said is that any answer exhibiting one or more of these three concepts provides evidence that the student has grasped at least some of the concepts that it was hoped the animated energy graph would support. Conversely, any answer exhibiting a lack of awareness of these concepts provides evidence that the student has not gained the understanding that it was hoped the energy graphs would support, and perhaps not attended to the graph at all. Other answers, such as those stated purely in terms of formalism, were not considered to provide sufficient evidence one way or the other. This produced a second dimension of mutually exclusive categories into which the answers to Q6 (version 3) and Q7 (all versions) could be sorted.

Table 33: Do student answers contain evidence for use of any of 3 concepts supported by the Animated Energy Bar Graph?

(1) PE and KE change in opposition to each other, when one goes up, the other goes down.
(2) TE reading for a system can change (without physical aspect of system having to change).
(3) PE and/or TE quantities can become negative.

Yes = 1 **(sufficient evidence)**	**No = 0** **(insufficient or no evidence)**
"If you have no *KE*, then you have *PE*, and vice versa." (1)	"Only if he were on the ground not moving"
"*TE* can be zero when he goes lower than the reference line." (2)	"*TE* can never change because energy is conserved."
"If *PE* becomes negative and is larger than *KE*" (3)	"$-PE > KE$"

Each of the answers in the left column gives evidence for use of one of the three concepts, as indicated in parentheses. I required elaboration beyond a formal expression or a yes/no answer to count the answer as having sufficient evidence for use of a concept.

4. Evidence for Student Understanding of a Key Relationship

A key relationship that it was hoped students would grasp was that the reference line can be moved and that this would change the readings one would obtain for the total energy and the potential energy of the system. An examination of the three answers in the left hand column in Table 33 above reveals that, though they exhibit awareness that the energy amounts can change, they do not exhibit awareness that the movement of the reference line could produce the change. The second answer mentions the reference line but gives no indication that it can be moved. However, there were student answers that did exhibit awareness of this relationship:

"*TE* can be negative if zero bar is placed higher than track, negative *PE* > positive *KE*."
"Depends on reference line, *PE* + *KE* < 0."

Even though some formalism is used in each of the above answers, neither is given purely in terms of formalism. Rather, they each give evidence of a fairly sophisticated and complete understanding of a central concept addressed by the lesson. This suggests a third dimension of mutually exclusive categories into which the answers to Q6 (version 3) and Q7 (all versions) can be sorted.

Table 34: Does student answer contain evidence for use of the key relationship supported by coordinated use of the two key features?

Key Relationship: *TE* and/or *PE* depend on position of the reference line, where the implication is that *both* position of the line and the energy amount can change.

Yes = 1 (sufficient evidence)	**No = 0 (insufficient or no evidence)**
"*TE* can be negative if zero bar is placed higher than track, neg *PE* > pos *KE*"	"*TE* can be negative when *h* is negative and object is below the ref line"
"Depends on ref line, *PE* + *KE* < 0"	"If skater never goes above the zero line"
"If zero point very high and *KE* low enough"	"If negative *PE* > *KE*, if trial is under the curve"

In each of the answers on the left, the implication is that *TE* and/or *PE* depend on the reference line and that this line can be changed. Although the third answer exhibits a possible lack of clarity about how *KE* figures into the scenario, it does exhibit understanding of the key relationship expressed above.

5. What Can Coding the Answers Along the Three Binary Dimensions Reveal?

Although none of this coding can reveal for sure where students gained these concepts, if there appears to be a pattern of differences between evidence for their use in the whole class and small group conditions, this would be interesting. Any such patterns can be examined in light of other patterns in the data. For instance, patterns in activity sheet data can be examined in light of patterns observed in pre-post data and in videotape data. Because different analytical methods have been used to analyze activity sheet, pre-post, and videotape data, such comparisons will not be quantitative. Each kind of analysis yields a unique snapshot taken from a different cross-section of the data, and yields a different perspective of the learning experiences of these students. Considered together, these snapshots have the potential to build up a coherent, more multidimensional image of these data than any one analytical method alone.

Results for Projectile Motion Activity Sheet coding are presented in Table 35.

6. Selected Gravitational Potential Energy Activity Sheet Explanation Questions: Results

Table 35: Class averages for Question #7 (or Questions #6 and #7)

Scoring **Evidence present: 1.00** **Evidence not present: 0.00**

Class	Teacher	N	Format	Activity sheet #6b (v3), #7	1) evidence for use of GPE reference line	2) evidence for use of concepts supported by bar graph	3) evidence for use of key relationship	
Yr 1 HP	Teacher A	20	WC	Could TE ever = 0?	0.10	0.05	0.05	
Yr 1 HP	Teacher A	18	SG	Could TE ever = 0?	0.00	0.00	0.00	
Yr 1 CP	Teacher B	11	WC	Could TE ever = 0?	0.36	0.27	0.18	
Yr 1 CP	Teacher B	13	SG	Could TE ever = 0?	0.00	0.00	0.00	
Yr 1 AP	Teacher B	13	WC	Could TE ever = 0?	0.15	0.23	0.08	*
Yr 1 AP	Teacher B	18	SG	Could TE ever = 0?	0.33	0.44	0.22	
Yr 2 AP	Teacher B	21	WC	Could PE ever be <0? Could TE ever be <0?	0.95	1.00	0.95	
Yr 2 AP	Teacher B	21	SG	Could PE ever be <0? Could TE ever be <0?	0.81	0.95	0.48	

*Teacher inadvertently skipped Q7 during whole class discussion

Even though the data set is not large enough to yield a statistical result, a trend can be seen: In every comparison in which the teacher led whole class discussion about Question #7, the students in the whole class discussion format appeared to outperform the students in the small group discussion format along each one of the three dimensions on the right. They showed more evidence for using the GPE reference line in their reasoning, more evidence for using the semi-quantitative relationships between different kinds of energy in their reasoning, and more evidence for using the relationship between the position of the GPE reference line and the amounts of *PE* and *TE*. (Because the same data were scored along all three dimensions, the results are not added across the dimensions.) These results can help inform discussion about Research Question #5 in Chapter VIII.

C. Activity Sheet Analysis: Projectile Motion Lessons

1. Selected Projectile Motion Activity Sheet Explanation Questions: Coding Criteria

The key features and relationships identified in the Projectile Motion Animations were the red arrows and how they changed with time in the Vectors Animation (Video Clip 1), the horizontal lines and variable spacing between them in Lines Animation I (Video Clip 2), and the vertical lines and constant spacing between them in Lines Animation II (Video Clip 3).

Four questions on the Day 2 Projectile Motion activity sheet (Appendix F) were selected for analysis for their potential to shed light on whether students actually understood what the features meant and whether they had grasped a central concept addressed by the features, that projectiles accelerate in the downward direction only. Fortunately, three of these questions directly asked for interpretation of the key features and the fourth directly targeted the concept of downward acceleration, so the coding was fairly straightforward. Table 36 below lists the coding criteria and scores applied to the student answers and Table 37 lists the results. All coding was done blind to whole class or small group condition.

Table 36: Scoring for Projectile Activity Sheet Questions

Vectors Animation Question 4a: Does *this animation show acceleration? If so, in what direction?*	
SCORE	
1	if answer said 'yes,' in the negative or downward direction
½	if answer said 'yes' but did not specify or incorrectly specified the direction
0	otherwise
Vectors Animation Question 4b: What*, in this animation, lets you know that?*	
SCORE	
1	if answer said the vertical arrow changes size and the tip of the arrow moves downward
½	if answer said the vertical arrow changes size
0	otherwise
Lines Animation 1 Question 2a: Which *component of the velocity do these lines give you information about?*	
SCORE	
1	if answer said "y" or "vertical" direction of velocity
½	if answer said "y" or "vertical" but indicated that the lines represented something other than velocity
0	otherwise
Lines Animation 2 Question 4a: Which *component of the velocity do these lines give you information about?*	
SCORE	
1	if answer said "x" or "horizontal" direction of velocity
½	if answer said "x" or "horizontal" but indicated that the lines represented something other than velocity
0	otherwise

The results of scoring are given in Table 37 below.

2. Selected Projectile Motion Activity Sheet Explanation Questions: Results

Table 37: Class averages for selected questions

Scoring **Correct: 1.00** **Partially correct: 0.50** **Incorrect: 0.00**

Class	Teacher	N	Format	Does the vectors animation show acceleration? If so, in what direction?	What in the vectors animation lets you know that there is or is not acceleration in this system?	Which component of velocity do the (horizontal) lines give you information about?	Which component of velocity do the (vertical) lines give you information about?	Avg
Yr 1 HP	Teacher A	21	WC	0.64	0.21	0.79	0.83	0.63
		25	SG	0.82	0.42	0.84	0.92	0.75
Yr 1 HP	Teacher C	17	WC	0.85	0.47	0.94	1.00	0.82
		19	WC	0.58	0.32	0.89	0.95	0.68
		19	SG	0.63	0.32	0.74	0.95	0.66
Yr 2 HP	Teacher A	18	WC	0.75	0.50	0.83	0.89	0.74
		24	SG	0.60	0.46	1.00	0.94	0.75
Yr 2 CP	Teacher A	14	WC	0.57	0.21	0.93	0.86	0.64
		10	SG	0.70	0.30	0.70	0.65	0.59
Yr 2 AP	Teacher B	22	WC	0.73	0.50	0.73	0.64	0.65
		23	SG	0.83	0.43	0.91	0.91	0.77

From analysis of student responses to selected questions on the Projectile Animations activity sheet (including written and drawn responses), it does appear that many of the students were able to recognize and use the visual features that were designed to give information about the presence and direction of acceleration in the system. However, no clear whole class/small group pattern emerged. In two of the matched sets, the whole class condition appears to have performed better; in two others, the small group appears to have performed better; and in one other, the results for the two conditions were essentially the same. These results can help inform discussion about Research Question #5 in Chapter VIII.

CHAPTER VII

VIDEOTAPE ANALYSIS AND CASE STUDY COMPARISONS

A. Videotape Analysis: Introduction

The first research question that guided the present study is the following:

1. Are students able to exhibit gains in conceptual reasoning on pre-post tests from lessons that incorporate certain interactive simulations and animations?

This question was addressed in Chapter V and the conclusion is that students did show gains. However, contrary to teachers' expectations, there appeared to be no pre-post advantage for the students in the small group condition. The remaining research questions will be examined in light of the results for Question 1. These questions are:

2. To what extent do students and teachers engage in discussion about key concepts while working with the simulations and animations?
3. To what extent do teachers and students respond to conceptual difficulties and misconceptions exhibited during work with the simulations and animations?
4. To what extent do teachers and students support the recognition, use, and interpretation of key visual features of the simulations and animations?
5. Do students recognize and use key visual features of the simulations and animations?

Pilot studies indicated that videotape analysis had the potential to identify evidence that could address Questions 2-4, and to a lesser extent, Question 5. (Results from activity sheet analysis in Chapter VI will also be used to help address Question 5.) The purpose of the present chapter, comprising a series of comparative case study analyses, is to see how the issues addressed by the research questions played out in the context of matched whole class and small group discussions, and in so doing, to examine several aspects of the classroom discussions deemed important for learning.

The key concepts and key features mentioned in the research questions are specific to each of the two lesson sequences. These foci were chosen as the result of a significant piloting process (see Chapter III and Chapter IV, Section G) in which I tried to identify concepts and features that appeared critical for learning and for which data could be collected. For instance, the features needed to be visible in the videotapes and the concepts had to be likely to provoke discussion.

The research questions and development of the coding criteria will be discussed in detail in the context of each of the two lesson sequences. Thick qualitative case study descriptions of each class will follow. The case studies within each matched set are then subjected to comparative analysis to compare the whole class and small group videotape results within that set and to relate those results to the pre-post and activity sheet results. These comparisons will attempt to shed light on the issues raised in the research questions as these issues played out in particular class settings. This chapter will not look for patterns across the entire study; that is saved for Chapter VIII.

B. Videotape Analysis: Gravitational Potential Energy Lesson Sequence

1. The Gravitational Potential Energy Lesson

The gravitational potential energy lessons were centered on "Energy Skate Park," a simulation from the PhET project at the University of Colorado (http://phet.colorado.edu/index.php). As described in Chapter III, the simulation has sections of track that can be rearranged and shaped, and several human and non-human characters with different masses that can skate on the track. It has a variety of visual tools to help students make sense of the animated imagery and focus on the abstract quantities under discussion: pie charts, bar graphs, a movable reference line to indicate the height chosen as the zero for gravitational potential energy, a ruler, animated line graphs (Figures 12-14). In addition, there is an option to have the skater leave behind a trail of dots, each of which can be clicked to obtain a read-out of quantities associated with the skater at that point in the path. The user can change the value of gravity by moving the skater and track to different planets or into space. Friction can be turned on or off and there are thrusters that can apply forces when in space. When selecting the simulation, the teachers had stated they liked the fact that it is manipulable; that its various visual charts change in real time; that the zero of potential energy (represented by the gravitational potential energy reference line, also referred to as the "GPE" or "zero potential energy" reference line) can be moved up or down; that the gravitating planet can be changed; and finally, that in their experience, students find it engaging and humorous.

As described in Chapter III, much of the lesson focused on the skater skating on the parabolic-shaped track shown in Figures 12-14. The teacher and students referred to this track configuration as a "half-pipe," though an actual half-pipe does not have this geometry. Objectives of the lesson were for students to begin to understand how potential and kinetic energy can change into each other, the relationship between gravitational potential and height, the arbitrary nature of the choice of potential energy reference line and how this choice affects the measured values of energy, and the relationship between gravitational potential energy and gravitating mass. The physicist's idea is that, with friction absent, the gravitational potential energy possessed by the skater at the top of the half-pipe is converted to kinetic energy as the skater accelerates toward the bottom of the half-pipe. The total of these energies remains constant unless the skater gains energy from or loses it to the environment. However, the value assigned to the potential energy—and, thus, the value calculated for the total energy—depends on the elevation the user has chosen to be at zero potential.

From past experience, the teachers had identified the idea of an arbitrary zero potential reference height (represented by the movable GPE reference line, Figure 14) as a particular stumbling block for their students, especially at the Honors and College Preparatory levels. A related conceptual difficulty was the idea of the existence of negative energy, especially negative total energy. Questions on the activity sheet (Appendix B) were designed to address these concepts directly. Preliminary analysis of four transcripts (Chapter III Section B) had indicated that discussion in connection with activity sheet Question 7 had provoked student questions in both small group and whole class discussions. In the analyses below, discussion in response to Question 7 will be of particular interest.

When videotaping in the classrooms, the camera was viewed as a proxy for the experience of an individual student. For whole class discussions, the camera took the position of a student in the back of the room. For small group discussions, the camera moved with the students as they broke into small groups and then assumed the viewpoint of a student within one of the small groups. Therefore, the camera was focused on only one small group discussion at a time and taped what an individual student might have seen and heard. The coded transcripts in either condition, then, can be thought of as reflecting attributes and features to which an individual student might have been exposed during work with the activity sheet.

2. Research Questions and Transcript Coding Criteria

For videotape analysis, I used elements of the constant comparative method to progress from writing moment-by-moment observations for substantial portions of videotape data, to identifying patterns in these observations, to defining and refining codes that could be used for selective coding across multiple videotapes. The refined codes developed to address Research Questions 2-5 are below.

a. Research Question 2: To What Extent do Students and Teachers Engage in Discussion About Key Concepts While Working with the PhET Energy Skate Park Simulation?

Code: Student or teacher mentions possibility of total energy of some system being zero.

Code: Student or teacher mentions possibility of some kind of energy value being negative.

Total time spent on such discussion was noted.

b. Research Question 3: To What Extent Do Teachers and Students Respond to Conceptual Difficulties and Misconceptions Exhibited during Work with the Energy Skate Park Simulation?

First, episodes were flagged where a student expressed frustration, confusion, or puzzlement in connection with ideas presented within the animation, the activity sheet, or the class discussion. Then videotape segments that fit either or both of the following codes were coded as evidence for support of student conceptual difficulties.

Code: Response to conceptual difficulty: Classroom activity following an episode flagged as exhibiting "evidence for conceptual difficulty" was considered a response if it bore some relationship to the expressed difficulty.

Code: Response to misconception: Classroom activity was considered a response to a misconception if it appeared to be an attempt to address a misconception. (There need be no videotape evidence for the actual presence of a misconception, only that the responder appeared to think it was a potential issue.) Response could be from teacher or students or both.

For either code, the amount of discussion time spent on addressing the apparent student conceptual difficulty was established. No attempt was made to separate these responses into teacher and student responses; many responses were in the nature of joint discussion with overlapping comments.

c. Research Question 4: To What Extent Do Teachers and Students Support the Use and/or Interpretation of Key Visual Features of the Energy Skate Park Simulation?

Visual features identified as key were the Gravitational Potential Energy (GPE) reference line and the animated Energy Bar Graph. (An additional feature was identified in one matched set of classes and will be discussed as a special circumstance within that case study comparison.) It was considered that a transcript segment showed evidence for student or teacher support of other students' recognition and use of key visual features if the following code applied.

Code: Student or teacher supports use and/or interpretation of a key visual feature or relationship in the simulation.*

*Here, by "interpretation of a feature," I mean the interpretation of its meaning, the development of some degree of understanding, as opposed to attaining rote knowledge of the feature or the ability to recreate a visual aspect through mimicry.

Student or teacher is observed doing one or more of the following:

1) **Selectively pointing out** some aspect of the key visual feature or relationship as part of an apparent attempt to help students use it or interpret its meaning**;**

2) **Giving a hint** to encourage use or interpretation of the meaning of the key visual feature or relationship;

3) **Gesturing in the air or over the display** to indicate the key visual feature or relationship as part of an apparent attempt to help students use it or interpret its meaning;

4) **Asking a question to prompt use or interpretation** of the meaning of the key visual feature or relationship;

5) **Suggesting a manipulation of the simulation** to assist with use or interpretation of the meaning of the key visual feature or relationship;

6) **Pointing out a limitation** to interpreting the meaning of the key visual feature or relationship.

Individual visual support 'moves' were identified and counted. Generally when any one of the six actions was undertaken in an attempt to provide visual support, it was considered a single visual support 'move.' If the teacher or student simultaneously engaged in more than one of these actions, such as *selectively pointing out a key visual feature* while simultaneously *asking a question to prompt students to interpret its meaning*, this was counted as a single move. In long support episodes, a pause for response or a shift in tactics (asking a different prompting question, for example) was considered to demarcate between moves. However, if the same move was repeated several times in a row, it was counted only once.

d. Research Question 5: Do Students Recognize and/or Use Key Visual Features of the Energy Skate Park Simulation?

A videotape segment was considered to show evidence for student recognition and/or use of one of the key features if the following code applied.

Code: Student refers, points to, or moves the Gravitational Potential Energy Reference Line or the Animated Bar Graph.

Even though student actions were of primary interest to address this research question, teacher actions were also coded in order to help reveal patterns in whole class and small group situations.

Code: Teacher refers, points to, or moves the Gravitational Potential Energy Reference Line or the Animated Bar Graph.

3. Case Study Comparisons of Videotape Coding Results: Gravitational Potential Energy

Eight Gravitational Potential Energy classes, comprising four matched sets, met the criteria for the study as described in Chapter IV Section B. Descriptions of videotape coding results will be organized around the research questions. For the most part, the research questions will be discussed in the order 3, 2, 4, 5 as this will result in more narrative clarity for the Gravitational Potential Energy lessons. Following the case study descriptions of the classes in each matched set, diagrams of transcript codes and tables of results will be used to facilitate a qualitative comparison of the matched classes.

a. Year One: Honors Physics (Teacher A)

Teacher A taught the Gravitational Potential Energy lesson sequence as a 2-period lesson on succeeding days. That length of time seemed about right for these two Honors Physics (mid-level) classes.

i. Whole Class Condition: Year One HP Teacher A

The teacher began the lesson by giving an introduction to the simulation in which he demonstrated its features and followed student suggestions for manipulating the Energy Skate Park track. After about five minutes, he turned to address the activity sheet questions in whole class discussion mode (Figure 15).

On the first day, there were a number of episodes of student difficulty and the teacher's efforts to address these did not appear to satisfy the students. Late in this period the students exhibited a behavior that the teacher believed was an indicator of boredom in his classes: they left the class one after the other to go to the restroom, keeping the restroom key in constant use. One student was observed reading his email on his laptop

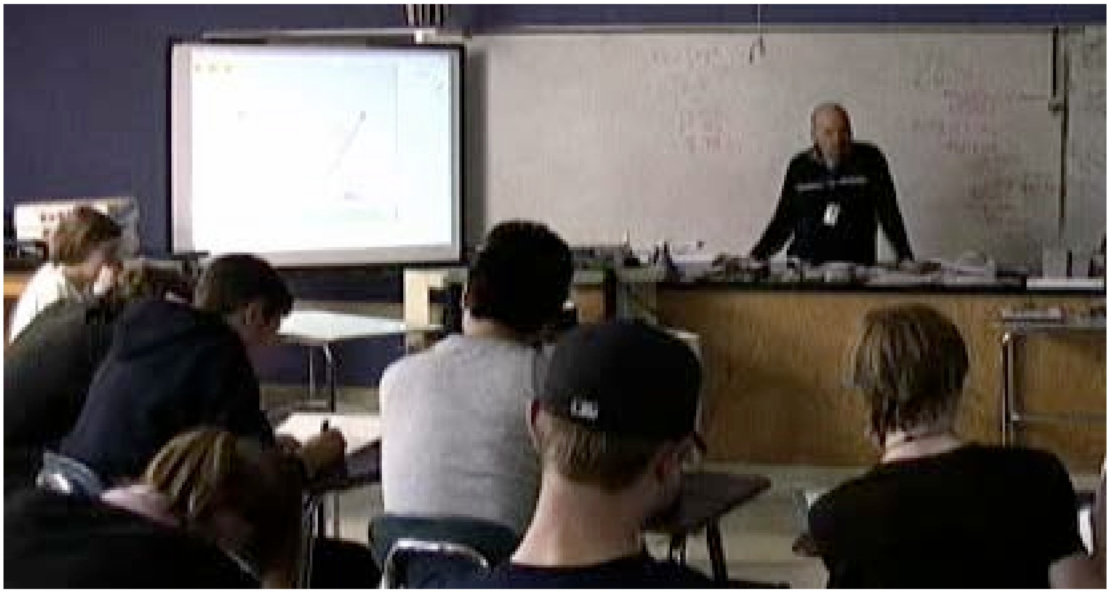


Figure 15: **Whole class working with *Energy Skate Park*, a PhET simulation.**

during the class. However, two incidents of student difficulty on the second day appeared to have a more successful outcome, with the students appearing satisfied. These incidents will be discussed further in the context of the research questions.

During this lesson sequence, the teacher began using a strategy that he was observed using increasingly the following year (in classes that were not included in this study because they failed to maintain comparable time on task). This strategy was to *set up the simulation to produce unexpected results* (producing what could be called a virtual *discrepant event*, Nussbaum & Novick, 1982) and, without comment, to wait for students to notice and react. This episode will be described in detail. Another large group discussion strategy this teacher occasionally used was to suggest that students turn to their neighbor and discuss an issue. In this class, he gave an implicit invitation by pausing to give time for such discussion to occur. These pauses were fairly short; the longest was 11 seconds. Frequently student-student discussion continued after these pauses had ended, while the teacher was turning to the next question, adjusting the simulation, or even after he had resumed talking.

Research Question #3: Response to conceptual difficulties and misconceptions. There were 11 episodes over the two days where students exhibited conceptual difficulty by expressing frustration, confusion, or puzzlement. There were 3 additional episodes where the teacher was observed addressing a misconception before the students appeared to be aware of any conceptual difficulty. Examples of these episodes, involving two different conceptual issues, are provided.

In the first episode, a student expressed puzzlement about thermal energy. During class discussion about an activity sheet question that asked how the skater's mass affected several different kinds of energy associated with the skater, the teacher had responded to a student question about thermal energy by turning on the Track Friction feature, even though he had not planned to address friction until later in the lesson. The animated Energy Bar Graph was on, and the puzzled student noticed that, as the virtual skater moved along the half pipe, although two of the bars in the animated bar graph rose and fell, the thermal energy bar rose but did not fall again.

201 S1: Would the thermal energy go down after a little while? 'Cause like, isn't it the case when you come to a stop, like-
202 S2 *(overlapping)*: Not in this program-
203 T: Yeah. So-
204 S3 *(Different student)*: Otherwise, every time- after, like, a week your tires would be on fire.
205 S1: Yeah.
206 T *(laughing)*: Good point.

In this excerpt, S2 responded to S1 by relating the behavior of the energy bar to the way the simulation was designed. S3 responded by proposing a thought experiment. He used an analogous system (car on road rather than skater on track) to investigate an implication of the theory that thermal energy would not leave such a system: moderate friction would then produce constantly increasing thermal energy. The result of the student's thought experiment was his prediction that the tires would catch fire. This result is not observed in every day life; his thought experiment produced logical evidence that tended to discount a theory of thermal energy as a quantity that could not leave the system. Much later in the class, the teacher explained that thermal energy would normally dissipate away from the skater by radiation even if there were nothing around to absorb it. But at this early point in the lesson, he merely made a comment (Line 206) that supported the student contribution without shutting down discussion. After the episode above, several students could be heard talking animatedly among themselves about the topic.

The remaining episodes that will be described are three that concerned the meaning of gravitational potential energy (GPE). The conceptual difficulty arose on Day 1, continued as the topic of discussion throughout the last five minutes of that day, arose again on Day 2, and continued for another three minutes. This issue first arose when the teacher reshaped the bottom of the track while the skater was skating along it. The animated energy bar graph was on and the PE bar rose, indicating that the skater had gained gravitational potential energy. The teacher realized that he had inadvertently raised the skater slightly as he was reshaping the track, delivering additional gravitational potential energy to the skater. The teacher started to move on to the next question, but a student wanted more explanation.

In the below, boldface indicates a *depictive gesture*, a gesture that appears to indicate an imaginary shape, location, or path in the air (Clement, 1994; Clement, Zietsman, and Monaghan, 2005). Such gestures will be described when they help convey the sense of the utterance.

251 S11: Wait, hold on. Does it actually change the energy or was it just because the potential energy went up because you lifted him up like ***[raises right hand above his head]*** fifty million yards off the ground?

252 T: The second thing you said. If I could somehow- like for example, if I were to *(clicks Pause)* Pause and then just do that- *(slowly drags bottom of track to the ground, then Unpauses; no noticeable change in height of skater or in size of energy pie chart).*

The discussion continued for a minute and then S11 tried again,

265 S11: What I'm getting confused of is, is the potential energy being changed by virtue of just, like….

The student continued with gestures but appeared to be having trouble articulating his question. The teacher responded to this expression of difficulty by showing the students the GPE reference line, considered in this study to be a key feature of the simulation. He explained that this was a point of reference and showed students how moving the reference line up and down changed the GPE readings on the charts. He moved the reference line above the skater and track and the GPE readings on the bar graph became negative. For a third time, S11 expressed puzzlement:

283 S11: Wait, hold on. So- then I have another question.

This time, without waiting for S11 to articulate his question, the teacher responded by giving an extended analogy with the temperature scale, explaining that different scales have different zero points. A change of temperature scale does not change what one feels when stepping outside. The teacher continued with a mini-lecture on this analogy, speaking for about 2 ½ minutes until the bell rang. Because he had no time for student feedback, it was not clear to what extent this analogy had helped the students.

The next day about 15 minutes into the discussion the question was raised again, this time by another student.

115 S6: That's something that really confuses me. Like, with the zero reference line, when you move, like, it just doesn't make sense, like the energy amount ***[hands move in front of his face as though containing something]*** due to that ***[moves finger up and down]*** reference line. Like how does that work?

It appears that the teacher's efforts had not cleared up the confusion; if anything, his use of the GPE line had increased the confusion. This was an issue observed in several classes; it did not appear to be at all intuitive to these students why the bars in the energy bar graph should change size when one moved an imaginary feature that had no effect on the physical set-up of the system.

The teacher responded to S6's comment by returning to the temperature scale analogy. A student interrupted by pointing out that, in other contexts, sea level can be used as a zero reference point for elevation. The teacher expanded on this point, talking about how a mountain does not look as high if the plain below it is already high. This entire response lasted about 3 minutes.

Additional episodes of student conceptual difficulty will be discussed in the context of other research questions below. In total, 14 minutes were spent on student difficulties, about 23% of this Honors Physics class discussion.

Research Question #2: Discussion about key concepts: Energy with zero or negative values. Concepts that had been identified in the pilot study as particularly difficult for students were the possibility of the *total energy of the system being equal to zero* or the possibility of *any kind of energy taking on negative values*.

Question 7 on the activity sheet directly asked about the possibility of the total energy of the system being equal to zero. This issue was reached 51 minutes into the 2-

day lesson sequence. (Compare with 23 minutes into the sequence for the matched Small Group discussion described below. The two classes had similar times on task but students and teacher did not use this time in the same ways.) The teacher introduced the second key concept, the possibility of any kind of energy having a negative value, during a mini-lecture near the end of whole class discussion about Question 7. He raised the issue again a few minutes later. No student explicitly mentioned negative energies during this discussion, although the teacher's comments appeared to elicit student questions about the nature of gravitational potential energy in general.

The transcript for the Question 7 discussion, about 2 ¾ minutes long, is included here in its entirety so that it can be compared later with the corresponding transcript segment from the matched Small Group discussion. The segment begins when the teacher read the question, which referred to the sum of the energies represented in the animated bar graph on screen: potential, kinetic, and thermal. Square brackets denote gestures and boldface denotes depictive gestures.

87 T: Now it's asking us, "Could the total energy be zero at some position?" How could we get a total energy of zero? *(Pause several seconds.)* Go ahead.
88 S5: Not in that situation. You only could if the person was on the ground, not moving.
89 T: Do they have to be on the ground not moving?
90 S5: Well, I said 'point of zero'.
91 T: OK.
92 S5: *(overlapping)* Whatever you are saying is zero height.
93 T: *(overlapping)* Alright, for example if I grabbed him, and said just, just chill right there. Right there, don't move. *(T grabs skater and puts him at the bottom of the half pipe, releases him. Skater doesn't move. Amusingly, he continues to change his facing at regular intervals as though still skating back and forth on the half pipe. The students watch quietly. T moves his hand in humorous imitation of the skater's movement and some students chuckle.)* Now if we look at the graph, *(turns on bar graph, points to bars which are barely visible right at the zero line)* we see that the energy is pretty much zero. Right? Maybe not exactly. Within a small amount. Please. *(Gesturing toward a student.)*

94 S4: Um, I think if you take the bottom point, then you move it down so it's more like a V, I think the potential energy decreases. I'm not exactly sure why, but-

95 T: Alright, so I think what (S4) is saying is that if we lower this point on the graph- *(drags bottom of track down; turns on something, the GPE line?)* do this- *(appears to lower line)* like that?

96 S4: Yeah.

97 T: The potential energy decreases? *(Appears to turn off GPE line.)*

98 S: *(off screen)* Yeah, but then you bring in the sides closer, I think.

99 *(Students confer. Teacher waits 11 seconds.)*

100 T: OK, so what (S5) is saying is that it depends on where the zero line is. So one of the ideas here is that *(points to ground on simulation)* there's the ground, and then there's- *(turns on GPE line)* the zero reference line. They're *(points to two different places in the air)* two different things. We can *(drags GPE line down to the ground from its prior position above the ground; though bottom of track has been dragged down to the ground, the ends of track remain up at their previous heights)* move the zero reference line around and we can put it on the ground. That's kind of a natural place for it. For example, in this room, where is the most natural place to call zero reference?

101 S: The floor.

102 T: Yeah, the floor, 'cause it's kind of difficult in this room for us to put things lower than that. *(Deliberately drops his pen on the floor.)* But we could also make it **[G]** higher up and say, **[G]** ooo now it's negative! Just as we could *(moves reference line 2/3 way up toward top of track)* move this reference line up and say, yeah now it's got negative energy and if I look at the graph, *(turns on bar graph; TE and PE bars are hanging down below the zero line; T points to them)* energies are negative. Just means above or- **[G]** Positive and negative now is not a direction, this is not a vector; it's a scalar quantity. But we can **[G]** arbitrarily make zero different places and say **[G]** more than zero, **[G]** less than zero. But it doesn't tell us **[G]** left and right or up and down or north and south. It's not a direction.

Interestingly, in this transcript excerpt, it can be seen that the teacher initially did not offer any new information. Rather, he used a tactic advocated by Minstrell (Van Zee & Minstrell, 1997); he rephrased S5's answer as a question, "Do they have to be on the ground not moving?" S5 then clarified her answer. The teacher, still without adding any new information, illustrated this student's suggestion with the simulation. Because S5's answer was only partially correct (holding the skater still at the zero height is not the only way she could have zero total energy), it might seem puzzling that the teacher did not make a stronger move. I suggest that the teacher's interaction with the student and affirmation of her comment helped keep the class engaged in active discussion; the discussion then continued rather than lapsing as a different student, S4, suggested that the half-pipe be changed so that its bottom point would rest upon the ground.

S4 actually took the discussion farther away from the point that the teacher was trying to make by equating the ground with the zero energy reference line, whereas S5 had appeared to understand that the zero height was relative. However, the teacher used S4's statement as an opportunity to illustrate the difference between the ground and the zero reference line. He drew the discussion about Question 7 to a close with a 60 second mini-lecture on negative energy. He illustrated the concept with the simulation and pointed to the animated bars on the energy bar graph, which were now below the zero line (Figure 16), indicating that both potential and total energy were negative.

The topic of negative energy arose again in a later episode when the teacher addressed yet another student's confusion about the zero reference line during discussion

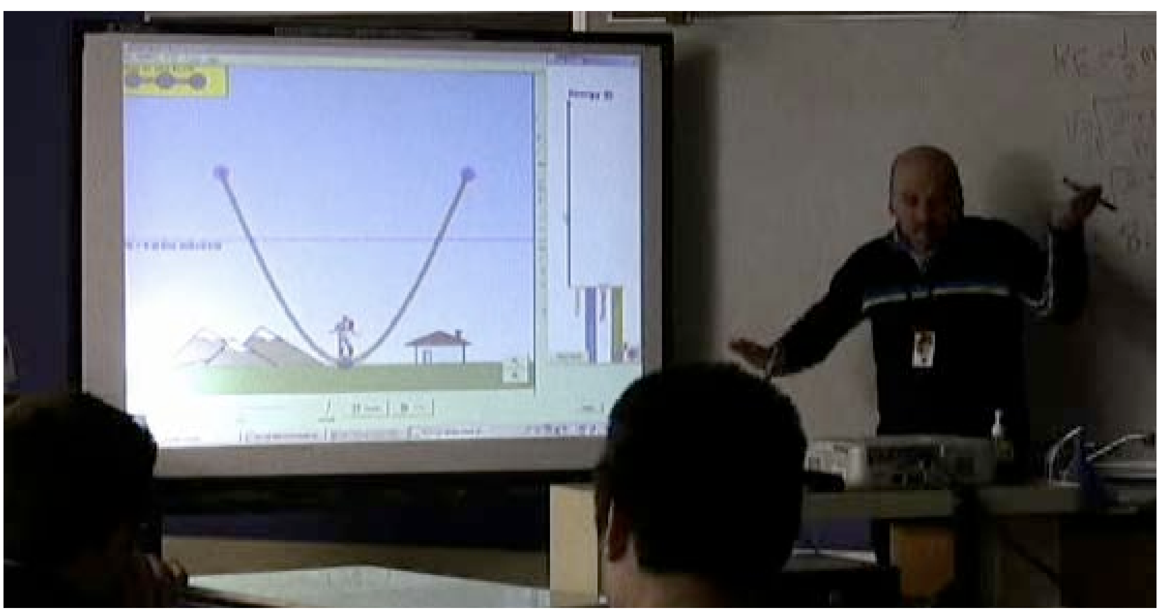


Figure 16: T: "(N)ow it's got negative energy and if I look at the graph, energies are negative."

about Question 8. To illustrate the arbitrariness of the potential energy readings, the teacher explained that in astronomy the zero reference is at infinity and all energy readings are negative. It was not clear whether this was helpful to the students.

Total discussion about these difficult key concepts during the whole class discussions of Day 1 and Day 2 was less than three minutes, about 5% of the discussion time.

Research Question #4: Support for key visual features: GPE reference line and energy bar graph. There was one student support episode and many teacher support episodes, four of which will be described. (See list of support moves, p. 120.)

Many of the teacher episodes involved *selectively pointing out aspects of the key features*, such as pointing to the tops of the bars in the energy bar graph, or holding a finger at their maximum heights while the bars oscillated.

208 T: So, the way that we change the skater's mass- so take a look, watch the energy gain. Watch what's happening on the *(points)* bar graph and watch what's happening *(points)* on the pie chart as we change the skater. *(Opens Choose Skater menu.)* Let's see, if I go to the bulldog skater *(changes skater)*, OK? What happened?

The teacher *selectively pointed out an aspect of the energy bar graph* to help students use it. He explicitly suggested that the students watch the energy gain as he changed another variable in the simulation, and then he asked them what had happened.

In the following, the teacher used two moves to support the students in interpreting and using the GPE reference line, another key visual feature.

267 T: … First of all, Reset and show this energy reference line- *(clicks RESET and turns GPE line back on)* -I can move this reference line up here. *(Moves GPE line from ground up to lowest point of track, above ground. Skater is skating.)* Now does that change the motion of the skater at all?

268 Ss: No, no.

269 T: But what it does do is, if I look at this bar graph and I see now the potential energy is zero at the bottom of the track, watch what happens if I move this down. *(As skater skates and T moves GPE line down to the ground, the TE bar rises within the bar graph and the PE bar continues to cycle up and down but no longer goes down to zero. All of this happens at once in different parts of the screen.)*

The teacher first followed instructions on the activity sheet to Reset and move the reference line. If a student or teacher merely followed written instructions, this was not counted as a visual support move. But then the teacher *asked a question to prompt the students to begin to interpret the meaning of the GPE line* by relating the movement of the line to what was happening in the simulation, “Now does that change the motion of the skater at all?” This was counted as one support move. Next, the teacher *selectively pointed out* what was happening to the bars in the bar graph as he moved the GPE line, “If I look at this bar graph…watch what happens if I move this down.” This suggestion appeared intended to *help students use* the bar graph and also to *help them interpret the relationship* between the bar graph and the GPE line. This was counted as a second support move.

In the following, the teacher used *gestures* to help students focus on the relationship between changes occurring to different bars in the bar graph. This was a relationship between relationships (second-order relationship) and these changes were yet to occur. Furthermore, the changes of interest would happen while the teacher was changing other aspects of the simulation. The square brackets with italics indicate gestures and boldface indicates a depictive gesture.

139 T: Rephrasing the question. If I switched to Jupiter, what will happen to ***[left hand above right, moves hands up and down, bringing them closer together and then farther apart]*** these bars? *[Points to bars on graph.]* Anybody wanna venture a guess before we try it?

The teacher had *gestured in the air to indicate a relationship between changes* that were about to occur in the bar graph.

The coding provides a very conservative estimate of the amount of teacher visual support present during this discussion. The teacher made many moves that did not fit the coding criteria, such as supporting the interpretation of visual elements not considered key features, pausing the simulation to prompt students to observe instantaneous visual relationships, or continually repeating an action in the simulation to give students a chance to see subtle but important changes in the animated charts and graphs. On one occasion, he made a continuous, repetitive gesture during student-student discussion when an issue that was not central to the discussion had the students sidetracked. The teacher appeared to be emphasizing that the important factor was the motion he was indicating with his hands rather than the issue the students were discussing.

One student visual support move was observed, on Day 2. The teacher had directed the students to observe how the energy was changing due to friction. There were many different kinds of movements occurring on screen at the same time: the skater cycled back and forth in diminishing arcs while in the animated energy graph, the TE decreased, the thermal energy increased, the KE cycled up and down between zero and diminishing maxima, and the PE cycled between a minimum value and diminishing maxima. Although many of the students appeared not to be clear on the function of the GPE reference line at this point, S1 appeared to have figured that out. He also appeared to have spotted an issue in the energy chart: the minimum of the PE was not zero and the changing maxima of the KE did not equal the instantaneous value of the decreasing TE.

111 S1: Where is the bottom bar set? *(Referring to GPE reference line.)*

112 T: Yeah, so that's an important question, 'cause when I hit RESET, it does something a little strange, which is that it puts the bottom of the track here *[points to lowest point on track, some distance above the ground]*, but the zero reference line is there *[points to the ground]*.

113 S1: Yeah, that's-

114 T: So maybe what I'll do is run this again but move the zero reference line to the bottom of the track. Okay?

Here, S1 has *asked a question to prompt use of the GPE reference line*, which the teacher appeared to be neglecting. In doing so, S1 has *selectively pointed out* the placement of the GPE line, crucial to interpreting the relationship between the GPE line and what was happening in the energy chart. He may also have been intending to point out a limitation of the simulation set-up: the teacher had neglected to move the GPE reference line back up to the bottom of the track after Resetting the simulation. This was counted as a single support move. This same issue arose in the matched small group discussion but was handled differently, as will be seen below.

There were 25 teacher support episodes and 1 student support episode, for a total of 26 visual support episodes, an average of about 25 episodes per hour.

Research Question #5: Recognition and/or use of key visual features: GPE reference line and energy bar graph. Although many episodes that provided visual support for the key features also involved use of the features or explicit reference to them, this was not always true. At times, a teacher or student merely gestured or gave a hint to support their interpretation or use. Conversely, some references to the features or use of them did not involve any moves identified as visual support moves. For instance, in the following Day 1 episode, the teacher made use of the energy bar graph and relied on the students to interpret its meaning:

199 S12: Is there no thermal energy 'cause there's no friction?

200 T: The answer here is yes, no thermal energy 'cause there is no friction. *(Plays simulation; skater skates along the half pipe. KE and PE bars in the energy bar graph swing up and down in opposition, the PE bar never reaching zero.)* This is what it looks like without the friction.

This was coded as *teacher uses key visual feature.* He modeled using the animated energy bar graph to obtain information about the thermal energy that was present in the simulation at a given time, although his move did not fit the criteria for a visual support move. A student example occurred a few moments later:

208 T: We are supposed to be … experimenting with how the skater's mass affects the different types of energies. … Let's see if I go to the bulldog skater *(changes skater)*, OK? What happened?

211 S3: Total energy is less. And there's no thermal.

When the teacher asked, "What happened?" he appeared to be referring to the energy bar graph. S3 appeared to be answering not from mathematical computation or from prior knowledge but from observing the changes in the bar graph and interpreting those changes. This episode was coded as *student uses key visual feature*.

The students were observed referring to or using the key features five times during the discussion. For comparison, the teacher was observed referring to or using the key features 26 times. Thus, students were exposed to some reference to or use of key visual features a total of 31 times, for an average of 30 episodes an hour.

ii. Small Group Condition: Year One HP Teacher A

As in the matched whole class condition, Teacher A taught the small group condition as a 2-period sequence on succeeding days. The teacher began the lesson by giving an introduction to the simulation in which he demonstrated its features, although he did not invite student suggestions for manipulating the simulation as he had done in the matched whole class discussion. Before the students broke up into small groups, the teacher introduced the activity sheet. He encouraged the students to finish the entire activity sheet the first day if they could, saying that not finishing it "wouldn't be the worst thing," and that they would have time the next day to finish if they needed it. The small group on camera took his request to heart, and paced themselves so that they could finish by the end of the first class period. Because many of the groups did not finish, the class spent an additional 25 minutes on task the next day.

The small group on camera had four students. (See Figure 1 in Chapter III, which discusses preliminary analysis of this discussion.) In order to get immediately to the numbered questions, they skipped the activity sheet instruction to spend 5 minutes exploring the simulation. They succeeded in finishing the activity sheet the first day and spent the entire discussion time the second day freely exploring the simulation rather than returning to activity sheet questions that had puzzled them. Although one student briefly requested that they return to an issue on the activity sheet, the other students in the group ignored his request. Their time on task with the activity sheet questions was 29 minutes compared to the hour that some of the other groups spent. Most of the episodes described below occurred on the first day, while the group was focused on the activity sheet.

Research Question #3: Response to conceptual difficulties and misconceptions. There were 8 episodes in this small group discussion during which students appeared to be experiencing conceptual difficulty. In 3 additional episodes, a student appeared to be responding to what he or she perceived as a misconception even though the students being addressed did not appear to be aware of experiencing any conceptual difficulty. On the first day of the discussion, these students responded to every expression of difficulty voiced by their group partners, although some responses were brief. On the second day while they were freely experimenting with the simulation, there were no clear expressions of conceptual difficulty although one of the students hinted at a possible difficulty when he requested a return to one of the activity sheet questions. The rest of the group did not respond to this request.

Several of the Day 1 episodes occurred when the students did not realize they were supposed to move the GPE line up to the bottom of the track before each investigation and, consequently, the results of running the simulation were not what they had expected. Other episodes occurred when students tried to apply a rule about the conservation of total energy in situations where it did not apply. Four of the Day 1 episodes will be described below as an introduction to the kinds of exchanges that were observed in this and other small group discussions observed in the study.

One episode occurred when two of the students appeared to be reasoning from a misconception about the nature of the total energy of a system. In response to a question on the activity sheet about where on the track different kinds of energy would be the greatest for the skater, S2 and S3 suggested that the total energy was the greatest when the skater was between the top and the bottom of the track, when he possessed both potential and kinetic energy. S1 disagreed.

133 S1: Shouldn't the energy be the same all the way through? *(Pause, the students look at each other.)* Because in a closed system energy is conserved?
134 S3 *(grins):* Alright.
135 S2: Oooh, OK. *(laughs)*

S1 responded to an apparent misconception of S2 and S3 (that total energy in a closed system may vary) by asking a prompting question, pausing, and then stating the rule. His lab partners laughed and agreed. Although S1 was correct in applying the rule in this situation, the students later referred to this rule in a situation where it did not apply. (See their discussion about Question 10 below.) The brevity of this response to a conceptual difficulty appeared typical of many small group discussions.

In another brief episode, the animated energy bar graph showed the presence of thermal energy even though the students had turned off track friction and then re-started the skater. This initially confused the students.

511 S1: Wait, why is there still thermal energy? *(Takes the mouse from S2.)* Hold on.
512 S4: I think there always will be a little bit.
513 S1: No, there shouldn't be. Why? *(Manipulates the simulation.)* Oh, when it hits, maybe?
514 S2: If there's any vibration.
515 S3: Yeah, when it hits, yeah.
516 S1: 'Cause when you *(inaudible)* from-
517 S3: Yeah.

The students have figured out that the skater's thermal energy increased whenever he collided with another object. There were two notable features in this episode: 1) each of the four partners participated in constructing an explanation to address the difficulty; and 2) a student attempted to use the simulation to investigate his own question. This was one of the few times students were observed taking the initiative to use the simulation to investigate their own questions other than when directly prompted by the teacher.

Another episode ended without a solution. S3 became confused about whether they were interested in changes to the maximum or the minimum value of the potential energy when they adjusted other parameters in the simulation. When S2 answered that they were using the maximum values, S3 became concerned that he had answered an earlier question incorrectly.

593 S3: So didn't we do this one wrong, then? *(Turns back to Question 2.)* This one, we said five thousand potential, but- *(Picks up the mouse and while the skater is skating, switches from Moon Location to Earth Location. However, the energy graphs show different readings for the Earth Location than the*

students had obtained earlier.) Look, we said five thousand, but wouldn't it be- Whatever. You know, it doesn't matter.

S3 tried to use the simulation to address his question about Question 2 by switching back to the Earth Location. However, he didn't realize that in the simulation the skater arrives in a new scenario with whatever kinetic energy he already has at the moment of the switch. Therefore the skater arrived in the Earth Location with the kinetic energy he had picked up in the Moon Location, and the energy readings were not what S3 expected. Such virtual discrepant events can serve to stimulate investigation, but S3 simply moved on to the next question. It may be that he did not know how to investigate his question. However, it should also be noted that, of the four students, S3 appeared to have taken most to heart the teacher's request to finish the activity sheet by the end of the class period. He appeared to keep track of the group's pace and periodically encouraged them to stay focused. This can be seen explicitly in one additional episode.

The students had predicted that the skater would make it around a loop in the track if they started him at any point on the track higher than the top of the loop (Question 11). They tried this and the skater fell. Laudably, these students did not try to explain away the fall or to argue that the skater had actually made it as did some other small groups, but accepted that their predictions had not been met and tried to figure out why.

643 S3: He doesn't make it.
644 S2 *(overlapping)*: Why? Why doesn't he make it?
645 S3: Doesn't have enough velocity.
646 *(S1 laughs)*
647 S4: Well, he only starts slightly above the top of the loop in general. And- *(points toward screen)*
648 S1: Yeah, so how high does he have to start then? *(Moves skater to a starting point at about 6m; the top of the track is at 6.5m.)*
649 S4: Pretty much at the top.

650 S2: If you start at the top, he can go.
651 *(S1 releases the skater from 6m and he makes it around the loop with only a very small skip in his path.)*
652 S3: Wow.
653 S1: So we have to start him all the way at the very top? *(Positions skater at the very highest point of the track, 6.5m.)*
654 S3: Yeah.
655 S2 *(overlapping)*: Yeah.
656 *(Skater makes it around the loop smoothly.)*
657 S3: It just barely does it, too.
658 S2: Yeah.
659 S4: I think for a loop it's just because- *(pause)* the potential energy is so high at the very top of the loop. *(Inaudible. Gesture but meaning is not clear.)*
660 S3 *(glancing at clock)*: Alright, we have like, five minutes. We could just finish this up. "Turn on the energy Pie Chart and Bar Graph."
661 S1: Wait, so what's the answer for b? *(S1 looks at S2's activity sheet as S2 writes.)* No, he falls. *(Writes)* "He- falls." *(Writes)* "Don't know."
662 *(S3 picks up the mouse and begins the next question.)*

S1 wrote “Don’t know” in response to the activity sheet request to explain the result, and read his answer aloud. It appears that he, at least, was not satisfied when S3 moved the group on.

In this group, response to conceptual difficulties was met with reasoning, explanations, and experimenting with the simulations. Even in the episode where the student response involved simply referring to a rule, the rule was accompanied by a prompting question. There is some indication that at least two of these episodes were shortened by a concern for the time; however, very brief responses to student conceptual difficulties were frequently observed in small groups in this study. In this group, response to misconceptions and/or conceptual difficulties totaled 3 minutes 21 seconds, or about 3% of this group’s discussion time.

Research Question #2: Discussion about key concepts: Energy with zero or negative values. The possibility that gravitational potential energy could take on a negative value was never mentioned in this group discussion (Day 1 or Day 2). The possibility that total energy could be zero was discussed during a single 45-second episode, in response to Activity Sheet Question 7. The transcript excerpt below begins at the point the students read Question 7 and ends when they moved to Question 8. It can be compared with the comparable excerpt from the matched whole class discussion above. During this segment, the simulation appeared much the way it had in Figure 1, a screenshot from a little earlier in this discussion, but the skater was no longer moving and the students had become focused on their activity sheets rather than on the computer screen.

422 S3: "Could the total energy be zero at some position? Explain." That would have to be that there is no kinetic and no potential and no thermal.
423 S1: Yeah-
424 S3: Which I don't think- Is that ever possible?
425 S1: No.
426 *(Students write for 9 sec.)*
427 S4: How about space?
428 S1: No.
429 *(S4 grins, makes an inaudible comment, then shrugs and shakes his head slightly in a gesture of good-natured defeat or exasperation.)*
430 S1: Absolute zero is only theoretical. So.
431 S4: Well, so is everything else in the world that was proven. So.
432 S1: Yeah, so it's only theoretical.
433 S3: Alright.
434 S4 *(to S1)*: *(inaudible)*
435 S1: Yeah, but it's still theoretical.

During this segment, although S3 appeared to have been wondering about the idea of zero total energy, when S1 said “No,” S3 did not question further. All four of the students began writing down their answers for Question 7. None of them questioned S3’s statement that, for the total energy to be zero, each kind of energy contributing to the total would also have to be zero. Rather, they became sidetracked by the question of whether thermal energy could ever actually equal zero. Although an interesting question, it was in some sense irrelevant to the question at hand because a negative potential energy, for instance, could still result in a total energy of zero. The idea of a negative value for energy did not arise in this group although the simulation is designed to show negative values for potential energy (with multiple visual tools) whenever the GPE reference line is raised above the bottom of the track. The activity sheet did not directly instruct the students to try the GPE line at positions higher than the bottom of the track. Though these students did, on occasion, use the simulation to try to investigate their questions, they did not do so for Question 7.

As mentioned in the preliminary analysis of the discussion about Question 7 for this class (Chapter III), of potential concern is the fact that the back and forth between the students did not develop into a substantial discussion of the concepts and they quickly moved on to the remaining problems on the activity sheet, having spent less than one minute talking and writing about this question. This can be compared with the 2 ¾ spent on this topic in the matched whole class discussion as described in the case study above.

Research Question #4: Support for key visual features: GPE reference line and energy bar graph. Two visual support episodes occurred when students *suggested a manipulation of the simulation to help with interpreting* what they were seeing. Other episodes involved students *selectively pointing out* the GPE reference line, changes within the animated energy bar graph, or relationships between the bar graph and other aspects of the visual display. (See list of support moves on p. 120.)

The following is an example of *suggesting a manipulation of the simulation to assist with use and interpretation of a relationship*:

458 S3 (*reading from the activity sheet*): "What happens to the maximum values of the four- " Can you start it over? *(S2 moves skater to the top of the track.)* All right, so thermal keeps going crazy, bumps up. Kinetic keeps getting smaller.

The students were engaged in joint reasoning and S3's move seemed intended to support the understanding of the group as a whole. (If he had been engaged in private exploration not shared with the rest of the group, this would not have been counted as a support move.)

Immediately after this, S1 *selectively pointed out a relationship* between a key visual feature and another feature in the simulation:

459 S2: So gravitational potential energy always present-
460 S3: Yep.
464 S1: It's only always present because our guy doesn't go down to the bottom.

S1 has selectively pointed out an important relationship between the non-zero minimum reached by the PE bar in the energy bar graph and the height above ground of the lowest point of the skater's track.

A few moments later, in another episode, there were three support moves in quick succession, each one selectively pointing out aspects of the animated bar graph to help with interpreting and using it. S1 had just referred to the amount of kinetic energy depicted in the bar graph by saying, "It goes back and forth."

481 S1: See, look. *(Replays simulation)*
482 S3: I know, but then it, uh- *(overlapping)* but then it decreases.
483 S2 *(overlapping)*: And it's not present when it changes direction.

First, S1, arguing that that the size of the bar was oscillating rather than uniformly decreasing, replayed the simulation to selectively point out this behavior. Then S3 pointed out the decreasing maxima, while S2 appeared to expand on S1's statement, pointing out at what points in the oscillation the bar reached zero.

Eight student visual support episodes were observed for an average of 16 student support episodes per hour.

The teacher was not observed making any visual support moves. He did not stop by this small group during their work on the activity sheet and had offered no support moves during his introduction to the simulation, reasoning that students could figure the features out themselves while working with the simulation in small groups.

Research Question #5: Recognition and/or use of key visual features: GPE reference line and energy bar graph. The students were observed making use of the key visual features at times when they weren't providing any visual support moves. For instance, shortly after the visual support episodes described above, the students began to describe what the thermal energy was doing.

490 S2: Thermal energy- (*Takes mouse and moves skater to top of track, releases him.)*
491 S1: Increasing in pulses.
492 S3: Keeps getting bigger over time. Pulses as-

Even though the students did not explicitly mention the energy bar graph, the only thing "pulsing" or "getting bigger" on the screen was the thermal bar in the energy bar graph. It appears that they were using the graph to obtain information about the behavior of the skater's thermal energy as he moved along the half pipe. Lines 491 and 492 were each coded as student use of a key visual feature.

The students referred to or were observed using the key features 21 times for an average of 43 times per hour. However, they missed instructions to turn on key features. At one point where the activity sheet instructed them to hide the GPE reference line, S3 began to read the instruction, "So what, uh- you need to hide the p- OK." The instruction may not have made sense to him because the group had not yet discovered the line; in any case, S3 did not finish reading the instruction and seemed unconcerned by it. This was an instance when the students could reasonably have been expected to refer to this key feature and did not do so. Their failure to attend to the feature, even after they had discovered it, may have played a role in some of their conceptual difficulties in connection with the activity sheet questions. Their failure to make use of the feature certainly limited their ability to employ the simulation to explore the key concepts of zero and negative energies.

Other comments. There were several episodes when one of the male students, S1, manually took the mouse from the female student, S2. The most extreme example follows.

149 *(S2 has grabbed the skater with the mouse rather than pausing the simulation, with the result that the readings on one of the animated energy graphs begin to fade.)*
150 S2: Wait, why is it leaving?
151 S1: Just keep it running and it will stay. *(S1 puts his hand on top of S2's hand as though to grab the mouse from her but then moves his hand away from hers.)*
152 S3: Yeah, keep it running, it's fine. Put it all the way up at the-
153 S1: It's fine, it doesn't have to be at the top.
154 *(S1 actually pulls S2's hand off the mouse and takes the mouse from her. S3 and S4 laugh. S1 replaces the skater at the top of the track and lets him go, drops the mouse.)*
155 S2: You guys are really annoying.

Although S2 clearly appeared to be annoyed, she continued to pick up the mouse and manipulate it throughout the class and did not protest when S1 took the mouse from her again later in the class.

On the second day, the group spent the class period building tracks and exploring preset track configurations. They were observed making predictions about what kinds of curves, slopes, and gaps would allow the skater to go along the track without leaving the track. Some of the students appeared to have an intuitive sense of what it would take to make it around a loop or to make it over a hill without leaving the track but they did not appear to relate these intuitions back to their unresolved questions of the previous day. In particular, they not only appeared to realize that the skater had to start higher than the loop to make it over, but they now appeared to have a sense of how *much* higher he had to start. There was not any discussion about why this was so, however.

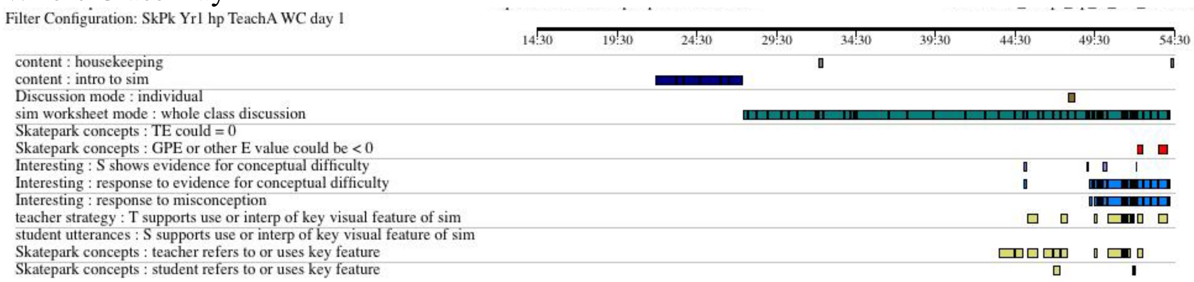
One of the students requested that they recreate a loop and try it with friction to see whether the skater would make it over. This was an extension of issues that had arisen for the group in connection with the activity sheet, but the other students ignored this request and the person controlling the mouse continued to explore preset configurations that did not involve loops.

The teacher did not stop by this small group during their work on the activity sheet. The group spent 29 minutes in discussion about the activity sheet questions.

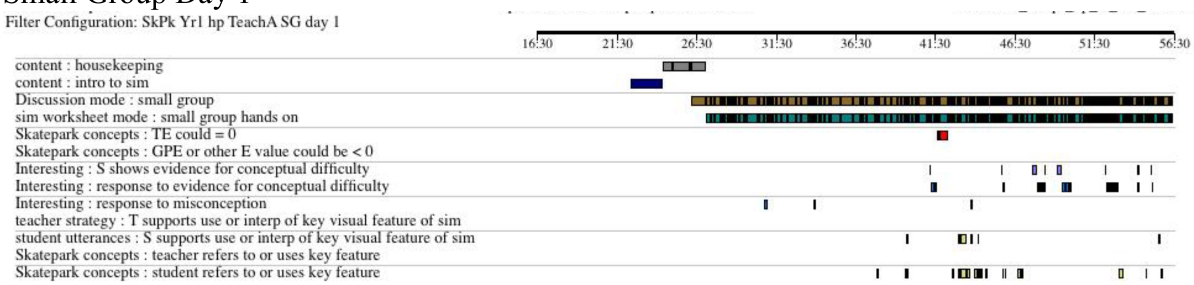
iii. Comparison: Year One HP Teacher A

In the videotape code maps (Figure 17), the transcript segments run chronologically from left to right, spanning the time when the students were working with the animations and animation activity sheets. Color blocks below each transcript segment denote the codes assigned to that segment; the code labels are listed on the left. The camera was used as a proxy for an individual student, staying in one place for whole class discussions, moving into a small group for small group discussions. Therefore, the codes can be considered to reflect what an individual student in that class might have experienced.

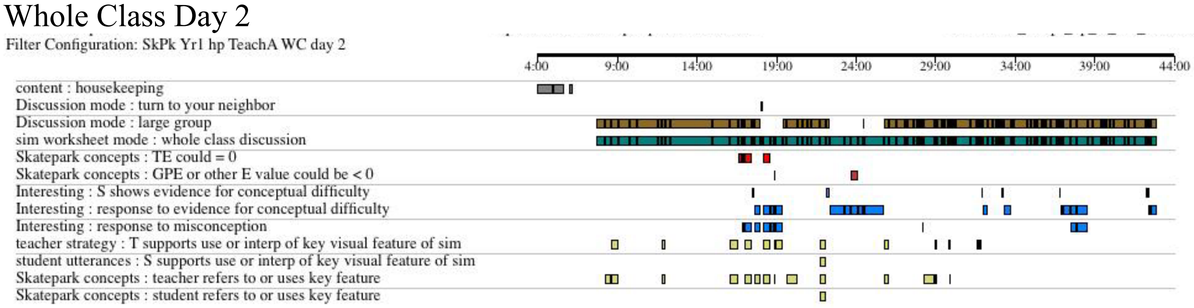
Whole Class Day 1



Small Group Day 1



Whole Class Day 2



Small Group Day 2

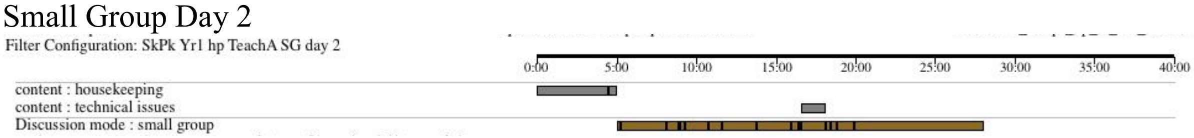


Figure 17: Videotape code maps: Year One HP Teacher A

Each timeline above represents 40 minutes of videotape, not all of which was taken up by classroom discussion. Small Group Day 2 had no episodes that matched coding criteria.

Table 38: Videotape coding results: Year One HP Teacher A
Time on task is given separately for the two days of the sequence.

	Whole Class Format	**Small Group Format**
Time provided for activity sheets (Hand out until pick up)	32 + 38 = 70 min	32 + 26 = 58 min
Time provided for simulations (including intro)	32 + 35 = 67 min*	32 + 25 = 57 min*
Time utilized by students on camera for activity sheet questions (Starting at Q1)	27 + 35 = 62 min	29 + 0 = 29 min**
Technical & other difficulties	1 ¼ min	1 ½ min
Length of taped discussion analyzed below	62 min 2 sec	29 min 14 sec
Research Q #2: Discussion about key concepts	Total length: **2 min 51 sec** Percentage of taped discussion: **5%**	Total length: **45 sec** Percentage of taped discussion: **3%**
Research Q #3: Response to conceptual difficulties and misconceptions	Episodes of difficulty: 11 Response length: 13 min 22 sec Response to misc w no prior evidence of diff: 3 Length: 41 sec Total: **14 min 3 sec** Percentage of discussion: **23%**	Episodes of difficulty: 8 Response length: 2 min 49 sec Response to misc w no prior evidence of diff: 3 Length: 32 sec Total: **3 min 21 sec** Percentage of discussion: **11%**
Research Q #4: Support for key visual features	Total support episodes: **26** Teacher: 25 Student: 1 Avg: **25 per hour**	Total support episodes: **8** Teacher: 0 Student: 8 Avg: **16 per hour**
Research Q #5: Recognition and/or use of key visual features	Total: **31** Teacher: 26 Student: 5 Avg: **30 per hour**	Total: **21** Teacher: 0 Student: 21 Avg: **43 per hour**

*Most small groups were finished so teacher called an end to the small group portion after 25 minutes. All students who were present on both days, and so included in the study, had finished their activity sheets. The matched whole class discussion went much longer than the teacher anticipated because of the length of the discussion following students' expressions of a persistent conceptual difficulty. (Interestingly, unlike in small group, two students in whole class had not filled out answers for the last problem on the activity sheet even though they had had longer time on task.)
**The small group on camera finished their activity sheet the first day. They spent the second day exploring the simulation but they did not return to the topics on the activity sheet, so time on task was considered to be zero.

Discussion. It can be seen from Table 38 above that, compared to the students in the small group on camera, the students in the whole class discussion were exposed to:

- **much longer discussion time**;
- slightly greater *percentage* of time spent on discussing *key concepts* though this was small;
- **more than double the *percentage* of discussion time responding to conceptual difficulties** (more than 4x the actual length of time)—although numbers of episodes in which students expressed difficulty did not differ greatly;
- greater rate of *support* for use of key features;
- but lesser rate of *references* to key features.

Probably the strongest different in the coding results was the difference in the amount of time spent on responding to student expressions of conceptual difficulties. It is possible that this had something to do with the significant difference on the pre-post tests in favor of the whole class condition, $t(37) = 2.22$, $p = 0.03$, $d = 0.71$. However, the results are richer than that.

For instance, the last bullet in the list above indicates that the students in the small group on camera referred to the key features more often than did students in the whole class discussion. This evidence addresses Research Question 5, about whether students used the key features. However, the activity sheet analysis (Table 35), which also addresses this question, revealed that *neither the small group on camera nor any of the other small groups in this class showed any written or drawn evidence* for actually having used either of the key features, or the relationship between them, in their reasoning about whether the total energy of a Skate Park system could ever equal zero. For the students in the whole class condition, there was evidence that at least one or two students in the class did use each of the two key features as well as the relationship between them. Clearly there is no statistical significance in such a result; two student outliers in the whole class condition could have produced this result. Overall, there does not appear to be a strong trend in the videotape and activity sheet results for this matched set for either the whole class or small group condition. In addition, even though the whole class condition had stronger gains for the short answer questions on the pre-post test, the small group format appeared to do better on the explanation questions, with 26% gains as opposed to 14% gains, although no statistical comparisons were run on these results.

This raises a question of what other factors might have influenced the pre-post results. There was a marked difference in discussion times that is important to consider and that can be compared in two ways. First, comparing the whole class discussion with the discussion time for the small group on camera, the whole class discussion was much longer. It can be seen from the code maps above that a large factor in this difference was that the small group used only about half of their allotted time. It will be seen in later comparisons that other small groups in the study also failed to utilize all of the time provided. Secondly, comparing the amount of time *that was made available* for the whole class discussion with the amount of time that was made available to the small groups—and that was fully utilized by some of the small groups in the class—there was still a ten-minute difference over the two days. (This was the largest difference allowed in the study; none of the other comparisons approached this difference.) The coding results cannot explain this. However, drawing on observation notes and subsequent interviews with the teacher, it can be seen that the reasons behind this difference illustrate some of the challenges in classroom management and how these challenges may play out differently in whole class and small group situations.

The teacher tried to keep time on task constant between the two classes, but this was the first time he had used this activity sheet and he vastly underestimated the time it would take the students to complete it. He taught the small group first on both days. He began the first day by encouraging the small group students to finish the twelve questions on the activity sheet on the first day if they could, although he also said that having the work spill over into a second day would not be "the worst thing." At least some students took this request seriously and paced themselves accordingly. In the whole class situation, on the other hand, when it became clear that students were experiencing an unexpected amount of conceptual difficulty concerning the nature of gravitational potential energy and the function of the GPE reference line, the teacher dropped his planned schedule and focused on this conceptual issue for the last 5 minutes of the discussion. This first day, the whole class discussion covered only Questions 1-4 and part of Question 5 out of 12 activity sheet questions, while the small group on camera finished all 12 questions.

On the second day in the small group situation, the teacher called a halt to the discussion when all of the students who had been present on both days had completed their activity sheets, which was after 25 minutes of small group discussion (in addition to the Day 1 discussion). The teacher fully expected to finish the matched whole class discussion in less than 25 minutes. However, when that discussion began, it became clear that students were still struggling, wanting to know how the total energy in the system could be changed merely due to repositioning an imaginary GPE reference line. The whole class discussion returned to this topic twice on the second day. The 10 minutes spent on this one conceptual issue can fully account for the 10-minute difference in time available for activity sheet discussion in the two classes.

One point that can be taken away from this is that, when unexpected conceptual difficulties arose, the small group on camera did not appear to feel able to give themselves permission to take time beyond what the teacher had requested them to spend. (Specific instances of this were revealed in student remarks quoted above.) In whole class, on the other hand, the teacher felt free, even impelled, to double the amount of time over what he had planned to spend on the lesson in order to address the difficulties in depth. Such situations would appear to present a dilemma for the teacher—how closely to stick to the lesson plan versus how much extra time to spend on unexpected student difficulties.

Even though this teacher chose to deviate from his lesson plan in this situation, it is not clear how much this extra time helped the whole class students. The teacher spent much of it on a rather abstract analogy using the relative nature of the temperature scale, when the students appeared to want more concrete explanations concerning how an imaginary line could have the power to change the total energy of a system. It was not the issue of relative values in general that confused them, but the idea that energy, in particular, could *be* a relative value.

Whether or not this extra time helped the class, because it was rather large (10 minutes) and not due to any intrinsic reason (e.g., having to do with what it takes, or normally could be expected to take, to orchestrate a whole class discussion vs. a set of small group discussions), the difference in pre-post gains will be presented in the conclusions section with an asterisk. The significant difference in favor of the whole class discussion could be a result of the teacher's underestimation of the amount of time it would take to complete the activity sheet, and a possible disproportionate effect this had on the two discussion formats regarding the information addressed by the short answer questions on the post-test. If the whole class discussion had been conducted before the small group discussions, it is likely the teacher would *not* have asked the small groups to try to finish in one day.

b. Year One: College Preparatory Physics (Teacher B)

Teacher B taught the Gravitational Potential Energy lesson sequence as a 1-period lesson. That length of time seemed a little short for these College Preparatory (CP) level classes, the least advanced physics level included in the study. This was the second semester that the teacher had used this simulation but the first time she had used it with this physics level. She had done some thinking about how to present it and had some idea of issues that were likely to prove problematic for these students.

i. Whole Class Condition: Year One CP Teacher B

About half way through the whole class discussion, the teacher opted to skip two questions on the activity sheet so that the class could get through the rest of the sheet by the end of the allotted time. This was successful, although she neglected one student question near the end of class in order to have time to make one final point she had planned to make.

On nine occasions, the teacher gave students the opportunity to turn to their neighbors and/or to write on their own. On six of these occasions, students were observed taking advantage of the opportunity to confer with their neighbors. The length of these episodes ranged from ½ to 1½ minutes. During these episodes, the teacher frequently walked around the room and answered individual questions. A total of around eight minutes was spent in these interludes.

Research Question #3: Response to conceptual difficulties and misconceptions. Students exhibited conceptual difficulties or puzzlement eight times during the class; all but the last of these episodes were followed by class discussion to address the difficulty. The last episode occurred as the teacher was wrapping up the discussion and she elected to ignore it in order to make the final point she had planned to make in the lesson.

Three episodes occurred early in the discussion when a student noticed bars on the energy bar graph drop below zero, prompting an unplanned discussion about the possibility of negative potential and total energy. This discussion about negative energy totaled about a minute and will be discussed below in connection with the key concepts.

In addition, the teacher appeared to address possible misconceptions a number of times before the students had exhibited any obvious conceptual difficulty. In total, 6 minutes were spent on student difficulties during this 42½ minutes college prep class discussion, or 14% of the discussion time.

Research Question #2: Discussion about key concepts: Energy with zero or negative values. A preliminary analysis of this discussion is in Chapter III, Section B.2.b. Key concepts were discussed twice, once for a minute and a quarter and again for three and a half minutes. These lively and entertaining excerpts will be given in their entirety so that they can be compared with the comparable discussion that occurred in the matched small group.

In this whole class discussion, the topic of negative energy arose before the topic of zero energy. The physicist's idea is that the gravitational potential energy possessed by the system when the skater is at the top of the track (referred to here as a "half-pipe") becomes converted to kinetic energy as the skater accelerates to the bottom and then becomes converted back to gravitational potential energy as the skater continues up the other side of the track. As the skater moves, friction causes some of the kinetic energy to be converted into thermal energy. The total of these energies remains constant unless energy is added to or subtracted from the system.

Before the following transcript excerpt began, the teacher had picked up the virtual skater and dropped him onto the track and onto the ground from various points so that the students could see the animated bars on the bar graph react. At one point the skater sailed off-screen and then fell below the potential energy reference line. Though the skater could not be seen, the potential energy bar in the chart fell into negative values. The following discussion ensued. (Students who could be seen on the videotape are referred to by number.)

249 S: Wait, he had negative potential energy, what?
250 S: Because he went below the line.
251 S: Oh, OK.
253 T *(more loudly)*: Yeah, yeah, this potential energy went negative. What's up with that, "Max"? What do you think?
254 S: He went below the line.
255 T: He fell below the line. So let me bring him back and catch him. If I move him down here, like I put him on the ground, he's got negative-
256 S: -Negative total- (*overlapping*)
257 T: -potential energy.- And negative *total* energy! That's interesting.
(*See Figure 2, p. 44. The potential and total energy bars are hanging down below zero.*)
258 S: And no thermal. Oh, you should throw him straight down to the ground and see what his thermal is.
259 (*Teacher does so.*)
260 S1: Whoa. Mad hot.

261 *(When skater hits ground, thermal bar shoots off the top of the chart.)*
262 S: Wait, is thermal, is more than total?
263 T (*repeats loudly*): Whoa, thermal is more than total?
264 S3: Because he has negative potential energy.
265 S: Oh snap.
266 S1: But you can’t really have negative potential energy in real life.
267 T: Well, it kind of depends. If you said the top- the roof of this building is my zero that I’m gonna define *(moves GPE line to roof of a building in the simulation)*, then when I'm on the ground it is negative. Not until I get myself up on the roof *(moves skater to the roof)* does it become zero. So, so it’s sort of semantics- I mean, it’s sort of like a definition, but yes, it can be negative. Usually, we choose the lowest point that we’re gonna get to, which if he stays in the half-pipe, is in the half-pipe. We usually choose that as our zero for potential energy.

This episode lasted 1¼ minutes, a minute of which concerned the existence of negative energy. A move that the teacher used here four times was to repeat certain student comments while adding emphasis, “-and negative *total* energy! That’s interesting.” She also manipulated the simulation in ways that appeared to arouse student interest; e.g., dropping the skater from various heights. This was an activity not suggested by the activity sheet, but one inspired by student questions. In the last line of the excerpt, the teacher provided an explanation for both the existence of an arbitrary zero height for potential energy and the existence of negative gravitational potential energy.

In spite of the above discussion, there was no guarantee that all the students understood or believed the teacher’s explanation. A short time later, she skipped Questions 5 and 6 to get to Question 7, about whether the total energy of the system could ever equal zero. The following 3 ½ minute transcript segment begins when the teacher read Question 7 and ends when she turned to Question 8. Note that even though the students had observed the skater with *negative* total energy earlier, this did not necessarily mean that they believed the skater could have *zero* total energy.

312 T: We're gonna zoom right over to Seven. And this is an interesting question, we kind of talked about this. "Could the total energy be zero at some point?"
313 Ss: No. No.
314 S: No.
315 S: 'Cause there is no such thing as [*inaudible*].
316 S: Yep.
317 S: And there is nothing there.
318 S: Well, maybe-
319 S: On the moon!
320 S: Even a rock has potential energy.
321 S: Not on the moon, on the Earth.
322 S: No, it can't.
323 S: Everything has energy.
324 S1: Oh yeah.
325 S: 'Cause it has the chance of moving.
326 S: Well, what if-
327 S: Earth is always moving.
328 S: Well what if, what if you just cement the [*inaudible*]?
329 (*Several overlapping comments from students, inaudible*.)
330 T: So, remember that potential energy reference line? Right now, the skater is sitting there- … *(Adjusts the simulation*) … The skater is sitting there and he's got lots of energy and it's always positive, right? His total energy is always positive. What could I do to maybe make his total energy be not so positive?
331 S: Start him at the line.
332 T: Start him what?
333 S1: Just move the line up. (*Referring to the zero potential energy line*.)
334 S: At the line.
335 S1: Move the line up to the top (*of the half-pipe*).
336 T: Move the line- the reference line? *(Closes control panel in the simulation.)*
337 S3: What would happen if you just put him at the very bottom?
338 *(T moves GPE ref line up; TE and PE bars go negative.)*
339 S: Told you. Wooo.
340 S1: Now move it down so the total gets to zero. No, up- Yeah, right there. *(S1 appears to be looking at the bar graph to see at what point the TE bar shrinks to zero.)* Oh I see, it has to be where he lands! *(Bar shrinks to zero when the line is placed at the top of the skater's arc.)*
341 T: Ahh!

(*The remaining two bars on the bar graph, for Kinetic and Potential Energy, are swinging up and down past zero, reaching zero together whenever the skater reaches the top of his arc*.)

342 S: *(Not clear, but sounds puzzled)*
343 S1: Where he stops for a second.
344 T: Where he's stopped? If you call the top of his rise, where he **[G]** stops for a second, cause when he **[G]** stops, his kinetic energy is- (*pause*) zero—and

you call **[G]** *that* the zero potential energy, then in a sense, total energy could be zero at some point. And what about if, uh, you just totally stopped him?
(*Teacher stops the skater at the bottom of the half-pipe. The TE and PE energy bars go negative and the KE bar disappears.)*

345 S: No.

346 S: No.

347 T: Yeah, let's put, tot- *(Moves GPE line to the bottom of the half pipe; all the energy bars disappear.)*

348 S: Yeah it's-

349 S: So it's all-

350 T: Yeah, I mean, he's not moving, right? He's not moving and he is down here at zero potential energy. He's got zero total energy. So yeah, what do you think? *(Some students begin conferring softly.)* It's a complicated question. There are many ways of answering it. If you just said yes or no, would that be a good way to answer a question like that?

351-353 *(Several students)*: No. No. No.

354 T: So you need to do a little explaining. So just take a minute or two, and see if you can write some kind of answer and explanation. You could say yes or you could say no, but you need to explain.

355 (*Students write for 50 sec*.)

The first thing to note is the large number of student-student exchanges here. Even though the teacher took a fairly strong hand in guiding the discussion, she was willing to take cues from students and to try their suggestions for operating the simulation. Occasionally she challenged the students with a question, "What could I do to maybe make his total energy be not so positive?" One student thought she knew how to get the total energy to zero and called out instructions that the teacher followed, resulting in the zero potential energy line being positioned at the top of the skater's arc. This did produce a total energy of zero, though the salient visual on the computer display was the kinetic and potential energy bars on the animated bar graph swinging wildly up and down in opposite directions. Eventually, the teacher stilled the skater at the bottom of the half-pipe, placed the zero potential energy line at that spot, and all the energy bars registered zero, but she suggested that this was not the complete answer. Finally, she prompted her students to write an answer that was more than a simple yes or no.

Total time spent on these concepts in this class discussion was 3 ¼ minutes on the idea that TE could equal zero and 1 minute on the concept of negative energy, for more than 4 minutes total, or 10% of the discussion time.

Research Question #4: Support for key visual features: GPE reference line and energy bar graph. The following extended excerpt gives an idea of the density of visual support strategies used by the teacher in parts of this lesson. Note that the act of identifying elements on the screen or of providing labels was not, by itself, sufficient to count as a visual support strategy. Also, one of the strategies described below was not counted in the summary table because it was being used to support the pie chart, not considered key in this lesson. In this excerpt, portions of transcript coded as <u>evidence for use of a support strategy</u> are underlined and a description of each strategy is inserted.

210 T: And I'm gonna turn on the bar graph and the pie graph. *(Turns them on.)* Let's take a look at what we've got here. The bar graph here shows energy. The green (*bar)* is kinetic energy, so <u>every time he goes to the top, it goes to zero *(follows bar up and down with the cursor)*, and he hits the middle and it's big.</u>

STRATEGY: Supports interpretation and use of a key feature (the energy bar graph) by selectively pointing out the movement of the tops of individual animated energy bars and pointing out their relationship to the skater's position on the track.

The blue (*bar*) is potential energy. It's big when it's up high and it's low when he gets to the bottom—but it's not zero. <u>Why isn't it zero?</u>

STRATEGY: Supports interpretation and use of a key feature (the energy bar graph) by asking a prompting question about the meaning of a relationship within the bar graph.

211 S: You didn't put the line back.

212 T: I didn't move the line back. So let's take care of, we'll take care of that in a second. Oh, I'll take care of it right now. *(Moves bar graph so that it does not obscure control panel.)* See, here is my potential energy reference line. *(Turns GPE line on; it is at ground level.)* <u>Watch what happens when I move it. *(Moves GPE line up to bottom of half-pipe.*</u> *Blue bar on bar graph now goes up and down from zero; green and blue bars move up and down in*

opposition to each other.) Now the potential energy goes to zero when he goes to the bottom.

STRATEGY: Supports interpretation and use of key features (the energy bar graph and the GPE reference line) by selectively pointing out a relationship between them.

Ahh, thermal energy, if there were any? If I turn friction on? You'd see that here (indicates blank space on bar graph), and total energy.

STRATEGY: Supports interpretation and use of a key feature (the energy bar graph) by selectively pointing out a blank place on the graph and indicating its meaning.

What's the first most striking thing about this graph?

STRATEGY: Supports interpretation and use of a key feature (the energy bar graph) by 1) asking a prompting question and 2) giving a hint.

213 S: Total never changes. *(Gold bar for TE doesn't change size.)*
214 T: Total never changes. That's called conservation of- energy, right? You can look, see the pie graph? *(Indicates with cursor.)* It's going from green to blue to green to blue? It kind of gives you that idea of energy [*facing away from the screen, moves hands back and forth in the air*] sloshing back and forth between kinetic energy and potential energy. And you can see that in the pie graph as well.

STRATEGY: Supports interpretation of the pie chart by gesturing to indicate a relationship between parts of the feature. NOT COUNTED.

If I pick the skater up, and drop him, first of all- watch what happens when I pick him up. *(She moves him up and down and the PE and TE bars move up and down. KE does not register.)*

STRATEGY: Supports interpretation and use of a key feature (the energy bar graph) by selectively pointing out relationship between movements of the bars and position of the skater in the air.

215 S: Total energy changes.
216 T: His total energy is changing. I'm doing some work on him. His total energy is changing.

There were a number of places where the teacher supported students' understanding and use of other visual elements within the simulation besides the key features. When working with the loop configuration, the teacher used many strategies to help students see the difference between the skater starting at *exactly* the same height as the top of the loop vs. *slightly higher* than the loop, vs. *enough higher* than the loop that he could make it around the loop without falling. In addition, she used several strategies to help students see that the slight hitch in the skater's movement as he went around the loop was not a glitch in the simulation but was the point at which a real live skater would leave the track and fall. This slight hitch was actually important evidence that the skater's initial potential energy had not been sufficient to provide enough kinetic energy at the top of the loop to keep the skater firmly against the track. The teacher provided these kinds of support an additional 19 times. Therefore the number given below provides a very conservative estimate of the frequency of visual support provided by the teacher during this whole class discussion.

In all, 29 teacher and 8 student episodes met the coding criteria for *support episodes for key visual features*. This was an average of about 41 teacher and 11 student episodes an hour, for a combined average of 52 support episodes an hour.

Research Question #5: Recognition and/or use of key visual features: GPE reference line and energy bar graph. As can be seen in the preceding section, key features were not necessarily explicitly mentioned during the support episodes and they may or may not have been actually used during those episodes. Therefore, it is of interest how often students referred to or used these features, whether or not this was in conjunction with the support episodes. Students were observed referring to or using the GPE reference line and /or the energy bar graph a total of 25 times during this class discussion, while, by comparison, the teacher was observed doing so a total of 33 times. This is a total of 58 times during the 43-minute discussion or an average of 82 times per hour that evidence for use of the features was observed. In the below, both student and teacher refer to a key feature, even though neither utterance meets the criteria for a visual support episode.

356 S: What's the name of that line again?
357 T: The reference line? Potential energy reference.

The frequency of references to, and use of, key features during this discussion will be contrasted with that during the matched small group discussion below.

ii. Small Group Condition: Year One CP Teacher B

Teacher B led the matched Small Group discussion class on the same day as the Whole Class discussion above. The 1-period lesson sequence seemed a little short for the small groups in this class; a lot of time in the small group on camera seemed to be taken up with logistics. The teacher began with a lengthy introduction to the simulation before the students broke into groups at computer stations. During the small group discussions, the teacher circulated the room answering questions and asking them. The small group that was joined by the video camera had two students.

Research Question #3: Response to conceptual difficulties and misconceptions. Students expressed slight puzzlement or surprise during two episodes. The first occurred early in the discussion, shortly after S1 had turned on friction for the first time. S2, who until then had seen the skater reach the same height on each side of the half pipe, expressed surprise that the skater was instead slowing to a stop at the bottom of the half pipe. S1 responded that this was because of friction. The two students then watched the simulation silently for several seconds before turning to other topics. The episode lasted about 13 seconds and involved little discussion.

The second episode occurred late in the discussion when the skater did not quite make it around the pre-set loop configuration but fell for part of the distance. This time both students expressed some surprise. The utterances that show evidence for student conceptual difficulty are underlined. The total discussion concerning this difficulty lasted 31 seconds.

538 *(Skater skips a part of the track on the way down and lands on a lower part of the loop.)*
539 S2: Whaa-
540 *(Skater continues to the end of the track.)*
541 S1: Oh, well, look, he kind of made it.
542 S2: No that doesn't count.
543 S1: Actually-
544 S2: Actually, he does not make it.
545 *(They write.)*
546 S1: I am not okay with that.
547 S2: Not make it because, because he didn't have enough potential energy.
548 S1: Because he didn't have enough kinetic energy?
549 S2: Potential.
550 S1: Kinetic energy is speed. He didn't have enough speed to go around.
551 S2: So, kinetic? *(writes)*

The emphasis in vocal tone on the videotape leads to an impression that the students were genuinely puzzled. However, once the students identified an important variable, they appeared satisfied and turned to the next problem; they did not engage in the rich discussion observed in the matched whole class discussion (see matched case study above) in response to the same skater behavior.

Two additional episodes were observed when student or teacher appeared to be responding to what they perceived as a misconception, while the students being addressed did not appear to be aware of any difficulty. The first episode occurred early in the discussion when S2 thought gravitational potential energy would be greatest at the bottom of the track.

180 S2: Okay, so the most potential- the most potential energy is, like, at the bottom. Right? *(Points to the bottom of the track, where the skater now sits at rest.)*
181 S1: No.
182 S2: Oh, the top.
183 S1 *(pointing to the top of the track)*: It's at the top.
184 S2: Oh yeah, the top.

S1 appeared to be responding to an apparent misconception about potential energy, but her response was a total of four words. Entire time spent on the issue was 10 seconds. The lack of discussion about *why* the correct answer was correct appeared to be typical of this small group.

In the second episode, the teacher appeared to respond to a misconception about total energy and suggested a manipulation for the simulation. However, the students, feeling confident of their answer, elected not to follow her suggestion. This episode is discussed further under Visual Support Episodes below. Response to misconceptions and/or conceptual difficulties totaled 54 seconds, or about 4% of the discussion time.

Research Question #2: Discussion about key concepts: Energy with zero or negative values. The possibility that total energy could equal zero in some situation was mentioned only once, when Question 7 was read. The possibility that gravitational potential energy could be negative was never mentioned during this discussion.

The transcript segment about the possibility that total energy could equal zero is given in its entirety. It begins when one of the students reads Question 7 and ends when the two students turn to Question 8.

432 S2: "Could the total energy be zero at the same position?" No, because you don't lose energy. You don't lose or gain energy.
433 S1: No, because energy is conserved.
434 S2: Yeah.
435 (*Students write for 13 sec.*)

The focus on this concept lasted 24 seconds, including the writing. Unlike in the matched whole class discussion, where focus on this topic had lasted more than eight times as long, this small group did not use the simulation to explore Question 7; this appeared to be typical of the small group discussions observed for this study. One hypothesis is that the students in such groups were in a "data collection mode," possibly their concept of what laboratory work is supposed to be. Their classroom experiences may have led them to view a "conceptual discussion mode" as something that occurs during whole class discussion rather than during lab. Another hypothesis is that, should these students implicitly have held a strong preconception that energy is a quantity akin to a substance and must be positive, it might have been unlikely to occur to them to explore other options or to test their ideas with the simulation, at least without external prompting.

Research Question #4: Support for key visual features: GPE reference line and energy bar graph. There were 2 student and 2 teacher support episodes that focused on key features. Both student support episodes were very simple: S2 twice reminded S1 to move the GPE reference line back up to the bottom of the track after a RESET (at points in the lesson where there were no explicit reminders to do this on the activity sheet). However, she did not explain why this should be done. These were—fairly generously—considered instances of *suggesting a manipulation to support use of the feature*.

The two teacher episodes occurred one after the other during one of the teacher's visits to the group and involved support of student use of the energy bar graph. When the students turned on friction and let the skater skate, the thermal bar on the bar graph began to grow:

457 S1: Ooo, they got a lot of thermal energy.
458 T: All of a sudden the thermal energy started to grow? *(Skater comes to rest at bottom of track.)* Did the total energy change?
459 S1 & S2: No.
460 T: Pick him up and start him over.
461 S2: No, you can't lose energy.
462 S1: Energy is conserved. Gosh, Ms. B
463 *(Teacher leaves laughing.)*

The first underlined utterance was considered an instance of *selectively pointing out a relationship involving a key feature*, in this case, the relationship between the current height of the total energy bar and its previous height. The second underlined utterance was considered an instance of *suggesting a manipulation to support using and interpreting the bar graph*. When the teacher suggested picking the skater up and starting over, it is plausible that she hoped that in repeating their actions, the students would more closely observe the changes in the energy bar heights and the relationships between them. It is also very possible that she could see their (incorrect) answers for the previous question written on their activity sheets, which indicated that these two students had some more thinking to do about the concept of total energy; they had been invoking the law of conservation of energy in situations where it did not apply. However, the students were confident that they understood the scenario and did not follow her suggestion. She elected not to pursue the issue further at that time.

There were additional episodes in which the focus of support was some visual aspect of the simulation that was not considered a key feature, one teacher and five student episodes. Although these episodes were not counted in the case study comparison, the presence of additional student episodes does reflect willingness on the part of these two students to help each other, even though this help tended to be in the form of very brief instructions unaccompanied by explanation.

Research Question #5: Recognition and/or use of key visual features: GPE reference line and energy bar graph. An estimate of how often the students actually used the features is provided by the fact that they referred to, pointed to, or manipulated them 9 times. Some of these were when S1 faithfully remembered to move the GPE reference line back up to the track after each RESET, although there was never any discussion about why this needed to be done. However, in many small groups, this instruction, near the beginning of the activity sheet, was either ignored or forgotten.

The teacher referred to the GPE line once and the energy bar graph once when stopping by the small group. Considering both student and teacher episodes, these features were observed in use or being referred to a total of 11 times during this small group discussion, for an average of 28 times per hour.

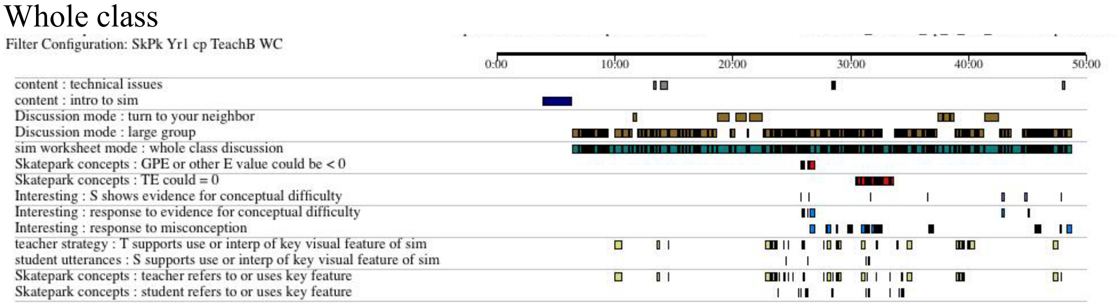
Other comments. The teacher stopped by the small group 4 times. She was present for a total of 1½ minutes out of the 24 minutes the group spent in discussion about the activity sheet. One of her visits was in response to the students' request that she come over because they were unsure about a calculation for the activity sheet. They stopped work on the activity sheet while waiting for her to arrive.

These two students spent 4 of their 24 minutes with the simulation and activity sheet either side-tracked with questions that were not pertinent, away from the table hunting for a calculator, or engaged in off-task behavior. Nonetheless, they had to skip only one question on the activity sheet in order to finish and were discussing returning to that question when the teacher called an end to the activity. This is consistent with other small groups, which often spent less time on the questions than students spent in whole class discussion, even when there was extra time at the end of the class period.

iii. Comparison: Year One CP Teacher B

In the videotape code maps (Figure 18), the transcript segments run chronologically from left to right, spanning the time when the students were working with the animations and animation activity sheets. Color blocks below each transcript segment denote the codes assigned to that segment; the codes are listed on the left.

Whole class



Small group

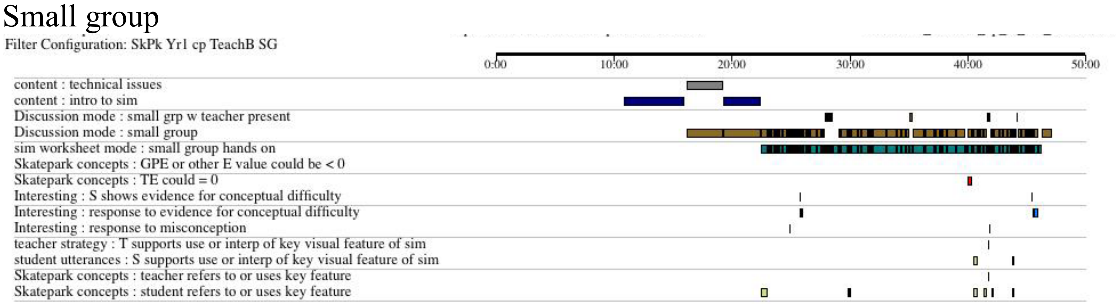


Figure 18: Videotape code maps: Year One CP Teacher B
Each timeline spans 50 minutes of videotape, not all of which was taken up by classroom discussion.

Table 39: Videotape coding results: Year One CP Teacher B

	Whole Class Format	Small Group Format
Time provided for activity sheets (Handed out until taken up)	49 min	42 min
Time provided for simulations (including intro)	45 min	36 min

Time utilized by students on camera for activity sheet questions (Starting at Q1)	42 ½ min	24 min
Technical & other difficulties	2 min 21 sec	4 min 52 sec
Length of taped discussion analyzed below	42 min 25 sec	23 min 54 sec
Research Q #2: Discussion about key concepts	Total length: **4 min 19 sec** Percentage of discussion: **10%**	Total length: **24 sec** Percentage of discussion: **2%**
Research Q #3: Response to conceptual difficulties and misconceptions	Episodes of difficulty: 8 Response length: 1 min 29 sec Response to misc w no prior evidence of diff: 33 Length: 4 min 40 sec Total: **6 min 9 sec** Percentage of discussion: **14%**	Episodes of difficulty: 2 Response length: 41 sec Response to misc w no prior evidence of diff: 2 Length: 13 sec Total: **54 sec** Percentage of discussion: **4%**
Research Q #4: Support for key visual features	Total support episodes: **37** Teacher: 29 Student: 8 Avg: **52 per hour**	Total support episodes: **4** Teacher: 2 Student: 2 Avg: **10 per hour**
Research Q #5: Recognition and/or use of key visual features	Total: **58** Teacher: 33 Student: 25 Avg: **82 per hour**	Total: **11** Teacher: 2 Student: 9 Avg: **28 per hour**

Discussion. It can be seen that, compared to the small group on camera, the whole class discussion had:

- **six times the *percentage* of discussion time** (10.2% vs. 1.7%) **spent on key concepts,** (more than 10x the amount of *actual* discussion time);
- **several times the *percentage* of discussion time spent on addressing conceptual difficulties and misconceptions;**
- **much greater frequency of support for using and interpreting key visual features;**
- **much greater frequency of recognizing and making use of key visual features.**

There was a 7 minute difference in total time available for work with the activity sheets in the small group and whole class conditions (first row in the chart), a small difference percentage wise and within the study's parameters for "matched classes," but still important to note as a possible contributing factor. In the small group class, the lesson had begun about 6 minutes later into the class period and ended a minute earlier to give students time to return to their desks.

However, this difference does not come close to explaining the difference in time on task for the two discussions. Several factors contributed to the difference. One was that the teacher spent longer introducing the simulation to the small group students before they broke up into groups, 5 minutes as compared with 2 ½ minutes in the whole class format. She reported later that she felt as though, for the whole class condition, she could wait to introduce many features until they naturally arose within the whole class work. For the small group condition on the other hand, she wanted to make sure that students were aware of the existence and location of certain important features before they broke up into their groups. For instance, she wanted to make sure that they knew to reset the position of the GPE reference line each time they reset the simulation. Her concern appeared to be warranted because this issue did arise in the small group on camera in spite of her efforts during the introduction.

Another factor was that once students moved into groups, most were not able to begin immediate work on the activity sheet. They had to sign into their computers, navigate to the simulation and wait for it to load. In the small group observed, this took 5 minutes. In addition, instructions on the activity sheet suggested that students play with the simulation before starting work. In whole class, this play was wrapped into the teacher's introduction to the simulation, while in the small group on camera, the students played with the simulation for an additional 3 minutes before the teacher asked the class to turn to the questions on the activity sheet. This resulted in Question 1 being addressed by most of the small groups almost 15 minutes later into the class period than in the whole class format.

On the other hand, by this point in the lesson, the small groups had had an opportunity to play and experiment hands-on with the simulation while the whole class students had not. In spite of the shorter time available for the activity sheet questions, the small group on camera finished all but one of the assigned questions and had time for off-task behavior and play. This is consistent with other matched class comparisons in which the small groups spent less time on the lesson, even in situations where plenty of time was available for them. Although the matched whole class discussion made it through all of the assigned questions, this came at the expense of the teacher ignoring a student question near the end of the discussion. Therefore, it appeared that time was a little short in both conditions.

The differences observed on the videotapes between the two conditions were striking. Many more student expressions of puzzlement occurred during the whole class discussion. When these were spoken softly, the teacher repeated them in a loud tone and facilitated a discussion about them. There was much more discussion in response to them, 1 minute 29 seconds as opposed to 41 seconds in the small group. For the important key concept concerning the possibility of total energy equaling zero at some position, **the small group spent 11 seconds in discussion and 13 seconds writing** while the **whole class discussion spent 140 seconds in discussion and 56 seconds writing**. In addition, the topic of negative energies arose in the whole class discussion as a result of a manipulation to the simulation by the teacher, and this resulted in an additional minute of discussion. This topic did not arise at all in the small group.

Even when adjusting for the shorter time on task by the small group, the differences were striking. The percentage of discussion time spent on the key concepts of zero and negative energies was many times greater in whole class discussion mode. Likewise, the percentage of discussion time spent on conceptual difficulties and misconceptions was many times greater in whole class discussion mode. The rate of visual support episodes was several times as great in the whole class discussion. Even ignoring support episodes by the teacher, *the students in the whole class discussion* engaged in episodes to support use and interpretation of key features at more than twice the rate of *the students in the small group*: 11 student support episodes per hour in whole class discussion as compared with 5 per hour in the small group. (The student rates were not computed separately in the table above.) Likewise, the rate of recognition and use of key visual features was several times as great in the whole class discussion. Although it could not be determined when the students in the whole class discussion were pointing to the key features, and they did not have the opportunity to manipulate them, *the students in the whole class discussion mentioned the key features* at a greater rate than the rate at which *the students in the small group mentioned, pointed to, or manipulated them*, 35 times per hour vs. 25 times per hour. (Again, the student rates were not computed separately in the table.)

The activity sheet analysis (Chapter VI) also supports the hypothesis that the students in the whole class condition used the features more. This analysis looked for evidence in student reasoning in response to questions about the presence in the system of negative energy values or a total energy of zero. There was evidence for use of the key features, the GPE reference line and the animated bar graph, by four of the whole class students, two of whom also considered the relationship between the features. There was no activity sheet evidence for use of these features by any of the students in any of the small groups.

It is interesting that greater evidence for use of the features by the whole class students, as judged by videotape analysis as well as by analysis of student writings and drawings on the activity sheet, did not appear to translate into increased ability to answer the two explanation questions on the post-test. These questions concerned a different aspect of the system, asking whether a marble would make it over a bump in a track or around a loop and why. The concepts targeted here involved the relationship between the starting point of the marble and the energy needed to make it over or around an obstacle; the acquisition of these concepts did not appear to depend on the use of the key features in the same way that the acquisition of a concept of negative energy did. The students in the small group condition appeared to outperform the whole class students on the two explanation questions (though no statistical analysis was done). Although the students in the two classes had similar pre-test scores on these questions—the whole class condition had a pre-instruction average of 23% and the small group had 20%—the students in the small group class outperformed the whole class students on the post-test 41% to 27%. The additional time and focus on the key features, the conceptual difficulties, and the key concepts in the whole class discussion also did not translate to a significant difference in the performance of the two classes on the pre-post short answer questions, $t(23) = 0.097$, $p = 0.92$, $d = 0.04$. It appears that the strengths of the groups in the two conditions were different; the evidence does not support concluding that there was an overall advantage for either condition.

That said, a potentially important observation is that the students in the small group on camera appeared to operate in a "data collection mode" rather than a "conceptual discussion mode," possibly reflecting their idea of what lab work is supposed to be. This appeared typical of a number of small groups observed in the study. One hypothesis is that these students may view a "conceptual discussion mode" as something that occurs during whole class discussion rather than during lab. If so, consistently successful work on conceptual issues in small groups may depend on changing the norms and attitudes of such students toward small group interactions.

c. Year One: Advanced Placement Physics (Teacher B)

Teacher B taught the Gravitational Potential Energy lesson sequence to matched sets of Advanced Placement (AP) classes during Years One and Two. This offers the opportunity to look at how a teacher shifted her strategies after experience with the simulation and activity sheet. Though such a comparison is not a central focus of the present study, the shift in strategies, as reflected in the four remaining case studies in this section, should be apparent and will be referred to where it helps explain the teacher's choices.

The Year One AP matched classes were taught during the first semester the teacher had used the simulation, the semester preceding her teaching of the (less advanced) CP classes described above. In this first semester, the pie chart was used in the activity sheet in places where the later activity sheets used the energy bar graph. Therefore, for purposes of this Year One AP comparison only, the pie chart is considered a key feature.

This teacher taught the lesson to all her classes as a one-period lesson. For narrative reasons, in this case study comparison, the research questions will be discussed in the order 5, 4, 3, 2.

i. Whole Class Condition: Year One AP Teacher B

The 1-period length seemed about right for this AP whole class discussion to cover the questions on the activity sheet, although one question was skipped (perhaps inadvertently), and the students did not have time to engage in the open ended activity suggested at the end of the sheet, to create and experiment with their own track configurations. For the last two minutes of the whole class discussion, the teacher used equations during what was essentially a mini-lecture to try to help students make sense of what they were seeing in the simulation.

Research Question #5: Recognition and/or use of key visual features: GPE reference line, energy bar graph, and energy pie chart. The students referred to the key features 5 times during the whole class discussion. Three of those times involved inquiring about the current location of the GPE reference line, a dotted line that marked the zero height for gravitational potential energy. The line could be made visible or invisible; when visible, it could be moved to change the height considered to be at zero gravitational potential energy. In the following example, the student query in Line 172 was unprompted. The reference line feature was turned off and so the current position of the line was not visible.

172 S: The reference line is still at the-
173 T: The reference line is right at the bottom, so what's the potential energy? What do you predict the potential energy on this?
174 S: Zero.

The student both *referred* to the reference line and *used* it to answer the teacher's question.

The teacher made frequent use of the key features in the simulation as evidenced by the 16 times she was observed referring to and/or manipulating them. Some episodes will appear in transcript excerpts below. Considering both student and teacher episodes, the key features were referred to, pointed to, or manipulated at an average rate of 31 times per hour.

Research Question #4: Support for key visual features: GPE reference line, energy bar graph, and energy pie chart. In this whole class discussion, the students were not observed engaging in visual support episodes, although it should be noted that many of the student-student exchanges were inaudible. However, there were a number of teacher episodes, where the teacher was observed supporting the students in making use of, or interpreting, important visual relationships involving the key features. Two notable examples of teacher support occurred during the following episode.

To understand this episode, it will be helpful to know what was taking place on the screen. Although the total energy of the system in the simulation remained constant as the skater moved, the way that energy was distributed between kinetic, potential, and thermal energies changed. When friction was turned on, the PE and KE energy bars in the animated bar graph oscillated up and down in opposition to each other, but the maximum heights they reached became smaller over time. The thermal bar "bumped up" slowly at an uneven rate, while the TE bar did not change. Because several kinds of visual changes were taking place at once, it could be a challenge for students to understand which kind of change the teacher wanted them to focus on. In the following episode, the teacher wanted the class to focus on what was happening to the *maxima* of the PE and KE bars, not on the changes occurring between each maximum.

The teacher had run the simulation twice with friction turned on and the energy bar graph and energy pie chart showing. As the episode began, she asked the students to answer the next question on the activity sheet. In the following, underlined text indicates utterances and gestures that were counted as visual support moves. Square brackets indicate gestures and boldface indicates depictive gestures.

229 T: So, what I want you to write down in that chart in Number Eight is, what happens to the maximum values of gravitational potential energy, kinetic energy, thermal energy and total energy. Just really quickly, what happens to the maximum value of total energy?

230 *(students conferring among themselves.)*

231 S: Wait, but is it asking us, this one, uh- cause, uh-

232 T: Okay, so I'll start. (*Indicates with mouse*) Here's $t = 0$. Over time, what's happening to kinetic energy?

233 S: Is it asking when it's at zero, though, like when the *(points toward screen)* bottom would be at zero potential energy?

234 *(students conferring among themselves intently)*

235 T: It's asking, there's a most- As this *(skater)* oscillates back and forth, each kind of energy except for total gets bigger and smaller, bigger and smaller. So what happens to the biggest it is, to the peak?

236 S: Oh.

The teacher has *selectively pointed out specific aspects of the energy bar graph to help students interpret its meaning.* At this point she could have stopped because the student appeared to understand the question. However, she chose to continue with additional support in the form of a series of prompting questions and gestures.

237 T: What does happen? Gravitational potential energy gets- *(pause)* over time is ***[hand held horizontally, slightly above head height, pats the air downward, apparently to indicate peak values of the GPE bar]*** getting?

238 S: Smaller.

239 T: ***[moves hand slightly downward]*** Smaller. Kinetic energy is getting?

240 S: Smaller.

241 T: But thermal energy is getting?

242 S: Bigger.

243	T:	Bigger.

In Line 237, the teacher simultaneously 1) *gestured in the air to indicate a visual relationship* and 2) *asked a question to prompt interpretation of the meaning of this relationship.* She then asked two more prompting questions related to the first. Lines 237, 239, and 241 were conservatively counted as a single visual support move; the additional prompting questions closely repeated the first one, essentially forming one multi-part question.

Incidentally, after this series of visual support episodes, in which the teacher had spent some time making sure the class understood what specific, and transitory, visual aspects of the energy bar graph to focus on, she shifted her attention to probing questions designed to get the students to think more deeply about the physics represented by the chart, such as, “Why didn’t my thermal energy become equal to my total energy?”

There were 19 episodes of visual support observed, all of them by the teacher, for an average of 28 visual support episodes per hour.

Research Question #3: Response to conceptual difficulties and misconceptions. There were three incidents of conceptual difficulty observed during this discussion. The first two were minor instances of puzzlement and were dealt with quickly. The third concerned an issue central to the lesson, what happened when the skater went around a pre-set track with a loop in it. The response to this third difficulty will be described.

The students had seen the skater make it around the loop in the track without falling when his starting point was somewhat higher than the top of the loop. However, when his starting point was only slightly higher than the top of the loop, he made it almost to the top but then fell. This surprised some of the students. S1, in particular, had difficulty reconciling this fact with another fact he already knew: a skater released into a valley on a frictionless track will rise to the same height on the opposite side. In the excerpt below, some students were addressing the difficulty among themselves while Students 1-4 were engaging with the teacher. Utterances from the second conversation have been omitted for clarity.

375 S1: Wait, is it because *r* is a circle? Cause you said that-
377 S1: *(sounding anxious, tense)* I thought you said that, if it was there at the same height, then they would make it.
378 S2: He might be-
380 T: He would make it back up to this height, but he'd lose speed.
381 S2: Because he still has a little bit of velocity-
382 *(A number of students are talking excitedly and animatedly among themselves.)*
384 S1: Wait, but if it was a perfect circle?
385 T: You do know, if you start *(slight chuckle)* at the same height-
387 S3: But if it's higher, then the velocity-
388 T: But is a *little* higher good enough?
389 S4: No, but we can figure out how much height.
391 T: In fact, did we figure out how much higher it needed to be?
392 *(Students still talking animatedly to each other.)*
393 S: It had to be less than two fifths of the original height.
394 *(With this comment, all the students get quieter.)*

At this point, the teacher and students have remembered that they had actually worked this problem out in an earlier class. In Line 393, an off camera student was rifling through his papers to find this work. (The answer is actually that the height of the loop must be less than 4/5 the starting height.) The teacher said it was time to end the discussion, but instead, she went to the white board and talked for two more minutes, using force diagrams and equations to explain that when the normal force became zero, the skater would fall.

The responses of the whole class discussion to the three episodes of student difficulty, including the teacher's impromptu explanation at the board, lasted 3 minutes 45 seconds, or about 9% of the discussion time.

Research Question #2: Discussion about key concepts: Energy with zero or negative values. The topic of negative energies arose twice. The first time there was discussion about whether gravitational potential energy, kinetic energy, and/or total energy could equal negative values. The students did not appear to have trouble with the idea that GPE and TE could be negative, even though the teacher never raised the GPE reference line high enough for students to actually observe that scenario. However, there was a short discussion about why KE can never equal negative values (because the speed would have to be an imaginary value for that to occur).

The second episode about negative energies was also brief:

181 S: Why is there potential energy? (At the bottom of the half pipe.)
182 T: Because I obviously didn't set the reference height quite at zero.
183 S: Oh, so it's negative potential energy.
184 T: Yes, just a teeny bit negative.

This episode reinforces the impression that the idea of negative energy was not problematic for these AP students.

Interestingly, the idea of zero energy *did* appear to be problematic. Question 7 asked whether the total energy could be zero at some position, but the teacher skipped this question, perhaps inadvertently. The topic of zero TE did not arise elsewhere during the discussion and that scenario never occurred onscreen. As was seen in the discussion about activity sheets in Chapter VI, almost half of the students in this class left Question 7 blank and none answered correctly. Even though these students did not appear to have trouble with the concept of *negative* energy, the only scenario for a total energy of *zero* that occurred to them was the scenario where each kind of energy equaled zero. This is an existence demonstration that it is not only in small group work where students may miss a central and important concept from the lesson.

ii. Small Group Condition: Year One AP Teacher B

In the matched whole class format described above, the Gravitational Potential Energy lesson had worked fairly well as a one-period sequence, although Question 7 was skipped and the students did not have a chance to experiment with their own track configurations. In the small group format, the students on camera seemed to need more time. When they reached Questions 7-10, they worked through them quickly because time was short, and they were not able to bring their discussion about Question 11 to a conclusion. These students did not always discover the visual aids recommended by the activity sheet but found their own or used logic in lieu of visuals. They also had repeated technical difficulties with the simulation. Although these difficulties did not appear to affect their ability to reason about the questions on the activity sheet, the students frequently commented on the difficulties and spent time trying to solve them.

The teacher circulated around the room and stopped by each small group a number of times; she visited this group 5 times during the 37 minutes of the small group discussion. The research questions will be discussed in the order 5, 4, 3, 2 for narrative reasons.

Research Question #5: Recognition and/or use of key visual features: GPE reference line, energy bar graph, and energy pie chart. Due to the design of the activity sheet, the GPE reference line, the energy bar graph, and the pie chart were all considered to be key features for this matched set of classes. However, in this small group class, the teacher did not bring up the energy bar graph during her introduction to the simulation and may not have brought up the GPE reference line either, though at one point she referred to "a line." She relied on the small groups to find these key features when instructed to by the activity sheet. During the 5-minute exploration that began the small group work, the four students on camera played with the track, the thrusters, different skaters, and different gravitational scenarios, but did not pull up *any* visual aids, key or non-key. (Non-key visual aids included features that had appeared, during pilot study observations, to encourage use of numerical calculation in addition to or instead of conceptual reasoning. These included the purple dots feature, which brought up numerical data for each point, several animated graphs with numbered axes, and a numbered grid.) Later, when the activity sheet instructed the students to use the GPE reference line, they were unable to find it. They also may not have found the energy bar graph when instructed to, though this was not clear on the videotape. At several points, they appeared to be seeking to adapt other features for purposes normally served by the key features. When it was not possible to tell what feature was being used, if the feature was being used in the manner intended for the key feature, this was counted in order to provide a generous estimate for the small group.

A transcript excerpt that reveals the trouble these students had in identifying key features begins with an attempt to identify and use the GPE reference line. The excerpt begins when the students read Question 3 on the activity sheet, which asked them to "Explore the potential energy reference line by clicking on it and moving the line around." Question 3a then asked how the position of the reference line affected height, PE, KE, and TE.

316 S1: *(reading)* "How does the position of the reference line affect each of the following?"
317 S4: What's the reference line? This thing? *(Points to border between ground and sky in the simulation. They don't have reference line turned on.)* Yeah, the reference line is the line, right? *(laughs)*

The students had interpreted a line in the background image of the simulation as the reference line. This was reasonable; had the reference line feature been turned on, this is where it would have been located. However, the line they were looking at was not movable. It represented the ground level, the default position for the reference line, but certainly not its only possible position. Because the reference line feature was not turned on, there was no way for the students to change the zero height of the potential energy; it was fixed at ground level.

Unable to change the reference height, S3 moved the track up and down instead. However, moving the track did not affect the “height” reading. Although the simulation appeared to be calculating the height from the ground, it was actually calculating height from the (currently invisible) reference line.

323 S1: (*reading*) “How does the- height change?” Like, well, height obviously increases or decreases.
324 TEACHER ARRIVES at the group. *(Students write. T watches silently.)*
325 S4: Potential energy increases, we know that.
326 S1: When you increase the height.
327 S4: Yeah.
328 T: Try it.
329 S4: Yeah. *(He takes the mouse.)*
330 S3: Open a graph-
331 T: Hit Reset, please. *(Referring to the RESET button in the simulation.)*
332 S4: Show Path? *(He is referring to the SHOW PATH button. He brings up a graph, closes it immediately, clicks RESET.)*
333 T: Uh, well-
334 S2: Not yet.
335 T: -turn on the grid. Show where the reference line is. See, you have to *(turn it on)* so you know where it is.
336 S4: All right, so it's, like, one point five.
337 T: *(with deliberate emphasis)* The *reference* line. No, no. The ref-er-ence line.

338 S3: Where is the reference line?
339 S4: What do you mean by that?
340 T: *(pointing to check box on the right of the screen):* Potential energy reference line?
341 S4: Oohh.

Not until the last line of this segment did a student actually recognize the feature. The teacher played an important supporting role; this transcript segment will be referred to again in the next section in terms of the visual support episodes.

It was not possible to tell for certain whether students ever used the animated energy bar graph. S2 repeatedly pulled up a particular feature and reasoned with it as though it were the bar graph, engaging in the kind of semi-quantitative reasoning about relationships that was a particular strength of the bar graph; such episodes were counted. For instance, when students referred to the feature and reasoned about relative changes in different kinds of energy, this was counted. However, at other times the students appeared to use quantitative data in lieu of reasoning about semi-quantitative relationships, referring to specific energy values and giving numerical answers rather than reasoning about relative changes; such episodes were not counted.

In all, the students were observed referring to or using the key features (or features presumed to be key) 15 times. The teacher was observed referring to or using the key features 5 times. This was equivalent to a rate of 37 episodes per hour.

Research Question #4: Support for key visual features: GPE reference line, energy bar graph, and energy pie chart. There were 4 student episodes that involved visual support for a key feature (or a feature presumed to be key). In addition to these episodes, there were other episodes where students appeared to be supporting features not considered key. Some of these episodes appeared to involve students seeking to adapt non-key features for purposes normally served by key features they had not been able to locate. The non-key episodes will be mentioned in order to try to give as complete account as possible of the amount of visual support engaged in by the students in this small group, although only episodes involving the key features (or features presumed to be key) will be counted and compared.

In the following excerpt, in which the feature in use was presumed to be key, two student support episodes were identified.

170 S3: Go to- Look at this.
171 *(S3 pulls up a different feature. Appears to be either the Energy vs. Position Graph or the Energy Bar Graph.)*
172 S2: Yeah, but we can use this to find the um-
173 *(Students look at screen.)*
174 S4: Potential energy.
175 S1: It's the potential energy that's most on top, but what is the actual number?
176 S3: But potential energy is making it so it goes to- all kinetic energy.

Although it is not possible to tell for sure which feature is up, it was referred to, supported, and used consistent with its having been the bar graph, a key feature. S3 appeared to be inviting his partners to use the feature to see the relationship between the change in PE and the change in KE. Two visual support moves were counted. In Line 170, he *suggested a manipulation to support using the feature*, showing his partners the feature and suggesting that they use it to address the question at hand. In Line 176, he *selectively pointed out a relationship represented by aspects of the feature*, the relationship between PE and KE, saying that all the potential energy "goes to," or becomes, kinetic energy. Two other student support moves later in the discussion concerned the GPE reference line.

A number of teacher support moves and one student move were identified in the following exchange. The first part of this exchange (Lines 335-341) was discussed in the previous section in connection with a student's *recognition* of the feature but is repeated here to highlight the *support* moves. In this excerpt, students were trying to reason about Question 3 without having been able to locate the moveable GPE reference line. Visual support moves are underlined.

335 T: -turn on the grid. Show where the reference line is. See, you have to *(turn it on)* so you know where it is.
336 S4: All right, so it's, like, one point five.
337 T: *(with deliberate emphasis)* The *reference* line. No, no. The ref-er-ence line.
338 S3: Where is the reference line?
339 S4: What do you mean by that?
340 T: *[pointing to check box on the right of the screen]* Potential energy reference line?
341 S4: Oohh.
342 T: So you know where it is. It will show it to you now.
343 S4: OK.
344 T: You can turn off the grid; it's kind of in the way.
345 S3: Yeah.
346 T: Yeah, OK. You see where it is? *(S4 nods.)* You can *move* that *(with emphasis, referring to the reference line)*.
347 S4: *(moving the reference line)* Oooh.
348 T: So, right now if you laid down-
349 S3: *(overlapping)* Put it right on the dot *(referring to large blue dot that marks the lowest point of the track)*.
350 T: -purple dots and asked what's the height at the bottom, you can figure that out. So, why don't you lay down some purple dots, pick a fixed spot on the bottom.
351 *(S4 apparently clicks SHOW PATH and lays down small purple dots.)*
352 S4: Yeah, I see this one.
353 T: You'll need to be sure that when you move the line, you have to lay down new purple dots.
354 S3: OK.

In this excerpt, the teacher used at least 6 distinct support moves to help the students locate and use the GPE reference line, and a student used yet another support move. (There are 7 underlined teacher utterances but the move in Line 340 was considered an extension of the move in Line 337.) The teacher moves ranged from *suggesting a basic manipulation* (to turn the feature on), to *selectively pointing it out* ("The *reference* line.

No, no. The ref-er-ence line”), to *suggesting a manipulation to assist with its use* (“you can *move* that”) to *suggesting manipulations of another feature to assist with its interpretation* (turn on the purple dots of the Show Path feature to help with interpreting the GPE line; whenever the GPE line is moved, lay down new purple dots). The student support move in Line 349 involved *suggesting a* (different) *manipulation to assist with its use*: S3 suggested to S4 that he place the GPE line accurately by putting it on top of a blue dot that marked the lowest point of the track.

Student visual support episodes for non-key features included support for use of a grid, possibly in lieu of the GPE reference line, and for selecting the purple dots in the Show Path feature, possibly in lieu of the energy bar graph. One episode involved trying to find a feature that would show the amount of heat. This last episode, near the end of the discussion, was interesting because it suggests that the students may not have ever found the energy bar graph—it would have shown them how the heat changed over time. These episodes suggest that there may have been as much or more student visual support for non-key features as for key features in this small group. Reporting only the totals for the key features would give the appearance of an advantage for the whole class condition that might not, in fact, have been present. Therefore, an explanatory note is included in the comparison table below concerning the student support episodes.

All 6 of the teacher visual support episodes identified in this transcript were present in the excerpt above. Together with the 4 student visual support episodes that supported the use of key features (or features presumed to be key), there was an average of 19 visual support episodes per hour observed for this small group discussion.

Research Question #3: Response to conceptual difficulties and misconceptions. Four episodes were coded as response to conceptual difficulty. There was an additional episode in which a student appeared to be trying to address the misconception of another when the student being addressed did not appear to be aware of any difficulty. In a final episode, the teacher anticipated a difficulty in the group and began to address it before it arose. Four of these six episodes will be described.

The responses to conceptual difficulty observed in this small group discussion were brief. The students appeared frequently to misunderstand or misinterpret each other's utterances and a number of episodes of difficulty arose from those. In such instances, it was often difficult to tell what part of the difficulty was conceptual and what part was due to poor communication. However, in the process of trying to clear up miscommunication, the students appeared to be developing clearer understandings about the concepts; the group normally did not move on to the next question until each member had had a chance to articulate their own understandings.

The following episode occurred after the teacher had shown the students how to turn on the GPE reference line. It began with S4 reading an activity sheet question that asked where to place the GPE line relative to the half-pipe track configuration. In this configuration, the lowest point of the track was above ground level. It was hoped that students would see that placing the reference line at the lowest point of the track rather than at ground level would "make sense" and would simplify the calculations. It was also hoped that this would help give students an intuitive feel for the arbitrary nature of the zero point for potential energy, a concept difficult for many high school students.

399 S4: *(reading)* "Is there a place where it would make the most sense to leave this line?"

400 S2: Right at the bottom.
401 S3: I would say, yeah, at the bottom of the curve.
402 S1: "Is there a place- " Yeah.
403 *(S3 runs simulation)*
404 S4: Then why is-
405 S1: So, the bottom would be like the equilibrium level.
406 S4 *(pointing to the screen):* Yeah, like the ground, the ground level. Why would, why would we-
407 S1: The reference line went at the bottom of the curve, not ground.
408 S4: Oh! *(He marks through something he had written.)*
409 S1: So then the bottom of the system is in equilibrium.

S4 had earlier moved the GPE line up to the bottom of the curve himself. His apparent confusion (Lines 404 and 406) may have been that he thought the line should go at ground level, or it may be that he thought his *partners* were saying that the line should go onto the ground, and it was this *mistaken* answer that did not make sense to him. Although S1 addressed the miscommunication, clarifying that they were talking about the bottom of the curve and not the bottom of the ground, she also twice explained that this was an 'equilibrium level.' It is not clear what she meant by this (because they were using a frictionless set-up, the skater would not ever come to rest at that level), but it does appear that *she interpreted S4's difficulty as at least partly conceptual* and so tried to explain *why* the line would go at the bottom of the curve. Therefore, her response was counted as a *response to* (an apparent) *conceptual difficulty.*

A moment later, S4 asked a question that made it clear that he was having difficulty with S1's explanation.

416 S4 *(to S1)*: Wait, why did you say that?
417 S1: What?
418 S4: "Bottom of the ramp so that- "
419 S1: So that way, like, the bottom of the system is at equilibrium-
420 S4: OK.
421 S1: -and the calculations are easier.

Once again it is not clear whether S4's difficulty was conceptual or was a communication issue due to the imprecision of S1's utterance about an 'equilibrium level.' It is also possible that S4 was merely trying to ascertain how his partners had phrased their answers before he wrote his answer down. However, S1 again answered *as though the issue were a conceptual one* and offered an additional explanation that was closer to the explanation the teacher was hoping for (i.e., that the calculations would be easier). This exchange was also counted as a *response to conceptual difficulty*.

There was one episode in which the student being addressed did not appear to be aware of any difficulty. S4 was looking at the screen to obtain the total energy printed there:

220 S4: Total is- thirty-nine.
221 S1: Total is what you had before; the number can't suddenly change.

S1 appeared to believe that S4 had a misconception about the conservation of energy of a system, reminding him "the number can't suddenly change," and her response was counted as a *response to a misconception.*

One final episode occurred when the teacher was present, much later in the discussion when the students had the loop configuration on screen. The students predicted that the skater would make it over the loop, "There's no reason he wouldn't." The teacher knew that other students had a tendency to discount the slight fall of the skater and to interpret what they were seeing as the skater "making it around." In addition, in this group the simulation was 'glitching,' starting and stopping and behaving strangely, and the falling of the skater could easily have been misinterpreted as another glitch. She headed off this possibility by saying, flatly, as soon as the simulation finished playing:

749 T: But he doesn't *(make it).* See if you can explain that.

She then suggested that they try starting the skater at a higher point. They did so, and this time he made it around the loop staying firmly in contact with the track the whole way. She left the group with a hint:

752 T: There is more velocity at the top, and then he does make it.

The researcher and teachers had observed that many students had such a strong misconception about gravitational potential energy that they tended to explain away the simulation results. In this instance, the teacher elected not to wait until students expressed doubt about the results but brought that issue to a close immediately, addressing the misperception before it arose in order to focus the students' attention on *why* the surprising result occurred. Once the students had accepted the result and expressed surprise and confusion in reaction to it, she did not stay with the group as they tried to resolve the conceptual issue. Rather, having focused them on the important conceptual question, she left them to reason about it on their own. The students reacted by doing something that was rarely observed in these classes—they took time to write down their incorrect prediction and their reasoning for it. They then returned to the activity sheet question, asking, "So why didn't he make it?" All four students offered suggestions and were still in the midst of animated discussion about the issue when class ended. The discussion when the teacher was present, and much of the discussion after she left, was counted as *response to misconception.*

Total discussion in response to conceptual difficulties was about one and a half minutes (1 min 36 sec) or 5% of the discussion time.

Research Question #2: Discussion about key concepts: Energy with zero or negative values. The possibility of *negative* energy was discussed seriously (as opposed to facetiously) just once, when the teacher was present and raised the question. Part of that discussion is recounted here. Earlier the students had joked about the possibility of negative friction but that was not counted.

483 T: What would happen to the total energy as you moved it? Could that go negative too? Kinetic and potential? Possible?
484 S3: Yeah.
485 S1: Isn't total energy always supposed to be *(inaudible)*?
486 S3: No. Because-
487 T: Does it depend on where that reference line is?
488 S3: Yeah.
489 S1: Yeah. Because it also depends on potential energy.

The students agreed that potential and total energy could be negative and the conversation turned to kinetic energy, which they decided could not ever be negative. The total length of discussion about negative energy values lasted 37 seconds.

The students discussed the possibility of a total energy of *zero* when answering Question 7. The discussion was short, beginning with two of the students reading the question simultaneously.

590 S4 & S1: "Could the total energy be zero at some position?"
591 S4: No, because-
592 S2: Well, not this position, but-
593 S3: Total energy?
594 S4: No, because-
595 S2: Total energy is gonna be the same.
596 S3: In this setup?
597 S2: In this setup. If we move the potential energy reference it wouldn't be, but in this setup.

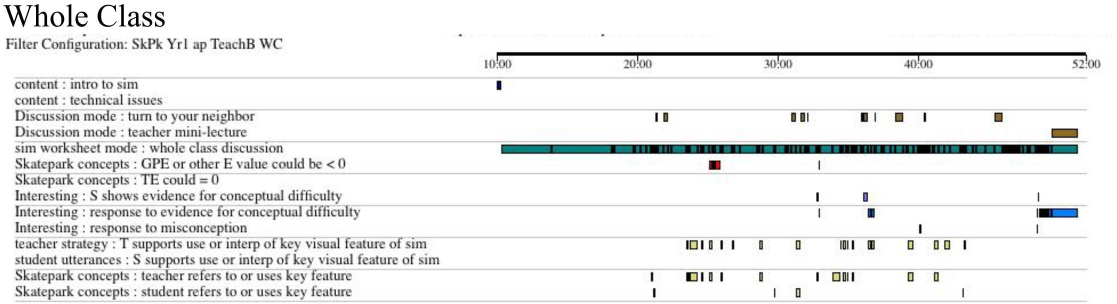
S2 understood that moving the GPE reference line could produce a TE of zero, and the other students appeared to have no difficulty with the concept. Even though S4 began by saying, "No," he may have been interpreting the question as applying specifically to their current setup, with the GPE reference line at the bottom of the track. Other parts of the discussion would tend to support the inference that these students were comfortable with the concept of zero TE. The episode lasted 23 seconds.

Total discussion about the key concepts of zero total energy or a negative value for any kind of energy was only about 60 seconds, or 3% of the discussion time, but this amount of time may have been sufficient for these advanced placement students.

iii. Comparison: Year One AP Teacher B

In the videotape code maps (Figure 19), the transcript segments run chronologically from left to right, spanning the time when the students were working with the animations and animation activity sheets. Color blocks below each transcript segment denote the codes assigned to that segment; the codes are listed on the left.

Whole Class



Small Group

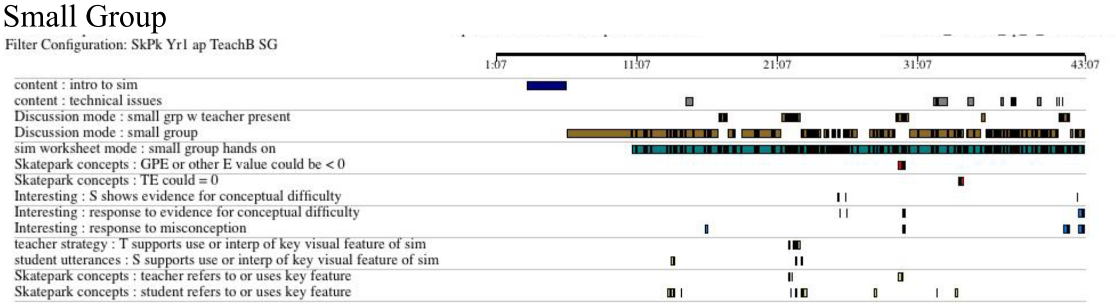


Figure 19: Videotape code maps: Year One AP Teacher B
(Each timeline above represents 42 minutes of videotape, not all of which was taken up by classroom discussion)

Table 40: Videotape coding results: Year One AP Teacher B

	Whole Class Format	**Small Group Format**
Time provided for activity sheets (Hand out until pick up)	44 min	37 min
Time provided for simulations (including intro)	44 min	40 min
Time utilized by students on camera for activity sheet questions (Starting at Q1)	41 min 6 sec	32 min 19 sec
Technical & other difficulties	0 min	3 min 20 sec*
Length of taped discussion analyzed below	41 min 6 sec	32 min 19 sec
Research Q #2: Discussion about key concepts	Total length: **55 sec** Percentage of discussion: **2%**	Total length: **59 sec** Percentage of discussion: **3%**
Research Q #3: Response to conceptual difficulties and misconceptions	Episodes of difficulty: 3 Response length: 3 min 31 sec Response to misc w no prior evidence of diff: 1 Length: 12 sec Total: **3 min 43 sec** Percentage of discussion: **9%**	Episodes of difficulty: 4 Response length: 55 sec Response to misc w no prior evidence of diff: 1 Length: 40 sec Total: **1 min 35 sec** Percentage of discussion: **5%**
Research Q #4: Support for key visual features**	Total support episodes: **19** Teacher: 19 Student: 0 Avg: **28 per hour** (25 per hour if the pie chart is not included)	Total support episodes: **10***** Teacher: 6*** Student: 4*** Avg: **19 per hour***** (There were no episodes involving the pie chart)
Research Q #5: Recognition and/or use of key visual features	Total: **21** Teacher: 16 Student: 5 Avg: **31 per hour** (28 per hour if the pie chart is not included)	Total: **18** Teacher: 5 Student: 13 Avg: **33 per hour** (30 per hour if the pie chart is not included)

* The difficulties did not stop students from using the simulation and so were not subtracted from time on task. The difficulties did however, constitute a distraction for the students.

** For these two classes, the pie chart was used in the activity sheet in places where the later activity sheets used the energy bar graph. Therefore, for purposes of this comparison, the pie chart was considered a key feature.
*** In this small group, students were initially unable to find the GPE reference line and perhaps were unable to find the energy bar graph. There were at least four student visual support episodes where the intent appeared to be to adapt non-key features for purposes the key features were intended to serve. At one point, the teacher supported those efforts before directing students' attention to the location of the key feature that was intended by the activity sheet. The starred figures represent only support for the key features, only about half of the total student visual support episodes noted for this class and one less than the teacher support episodes noted. In most of the comparisons in this study, the whole class discussions had noticeably more support episodes and a greater rate of such episodes than did the matched small group discussion, and these figures would appear to support that, but because of the difficulty in making a fair comparison in the present case, no such claim will be made here.

Discussion. As compared to the whole class condition, the small group on camera had:

- similar amount of discussion for the key concepts;
- similar rate of recognition and use of key features;
- only **about half the amount of time spent on addressing conceptual difficulties and misconceptions** although this was small in both classes (9% vs. 5% of discussion time)**.**

These discussions appear to have been more equitable in many respects than the discussions compared previously for the less advanced physics classes. In the present comparison, disregarding the results for "support of key features" where a fair comparison is problematic, the most notable difference between the two discussions is in the amount of time spent on addressing conceptual difficulties and misconceptions. The small group on camera ran out of time and was in the midst of discussing a conceptual difficulty when the class ended.

The small group did not discover one key feature until the teacher showed it to them about a third of the way into the discussion time. It is possible that they did not ever discover the other key feature, but if so, they were able to adapt a non-key feature for the intended purpose. Unfortunately, as described in the footnotes to Table 31, one result is that the number and rate of visual support episodes cannot be compared in an equitable fashion with those of the whole class discussion because the small group students supported each other in adapting and using non-key features to accomplish the tasks designed for the key features. However, one comparison can be made *within* the small group: the *teacher contributed more support episodes during her brief visits to the small group than the students contributed during their entire small group discussion*.

Another difficulty in comparing these two classes is that the teacher accidentally skipped Activity Sheet Question #7 in the whole class discussion. Student answers to this question, which asked whether the total energy of the system being explored (skater on a "Half Pipe," see Figures 12-14) could ever equal zero, normally played an important role in evaluating student use of the key features. Pilot studies had indicated that students who had not used the key features were not apt to discover the general answer to this question, that placing the GPE reference line at the top of a skater's arc would result in a negative potential energy that, at all positions, would equal the positive kinetic energy. But because the teacher skipped this question in whole class discussion, almost half of the students in that class left that question blank on their activity sheets. Of the 13 students who attempted to reason about this question on their own (out of 23 students), most did not arrive at the general answer. Only 5 students out of the 13 gave evidence in their written and drawn answers for using one or both key features and only one student in this class gave evidence for using the relationship between the features. Small group students did better; 18 out of 21 students attempted an answer; 11 of these gave evidence for using at least one of the key features; and 3 gave evidence for using both features and the relationship between them.

It is difficult to know how to interpret the above results. The students in whole class may have been at a disadvantage in their reasoning or they may simply have not have had time to think about Question 7 since the discussion did not pause to spend time on it. Even though videotape analysis can only compare the whole class discussion with the small group discussion joined by the camera, this analysis does span the entire length of the two discussions rather than just discussion about Question 7. Given the exigencies of these classes, this was probably the fairest comparison. Even though students in the small group may not have discovered all of the key features, if a feature was *used* consistently with its having been a key feature, this *use* could be counted. Analyzing the videotapes in this way revealed a similar amount of use in the on-camera discussions.

To summarize, at least one of the small groups had a disadvantage of not having discovered one of the key features and the whole class had a disadvantage of not discussing Question 7 on the activity sheet. The pre-post gains of the short answer and explanation questions are particularly valuable here to provide another snapshot of these classes.

There was a significant difference in pre-post gains in favor of the whole class condition: $t(42) = 2.37$, $p = 0.02$. (The pre-test averages of the two classes for the short answer questions were almost identical.) The effect size for this difference was medium at $d = 0.71$. In fact, the whole class had significant pre-post gains while the gains in the small group condition did not reach significance ($p < 0.001$, $d = 0.60$ and $p = 0.44$, $d =$

0.18 respectively). The difference in the Hake gains was even more dramatic at 42% and 8% for whole class and small group conditions, respectively.

The whole class group also did better on the pre-post explanation questions about whether a marble would make it to the end of a track with an obstacle and why; they had 24% gains while the students in all the small groups averaged only 2% gains on these questions. Therefore, even though Question 7 was skipped in whole class, it seems that something in the discussion did work for the whole class students. As mentioned, the small group on camera ran out of time and did not have an opportunity to complete their discussion of the loop problem, the problem most closely related to the discussion questions on the post-test. In the whole class condition, there was time for a rich discussion about this problem and the fact that the skater did not make it around. This was despite the fact that the whole class and small group reached this problem at almost exactly the same time into their respective discussions: the whole class about 33 minutes and the small group about 34 minutes into their discussions. However, as frequently seemed to happen in small group conditions, the small group on camera had begun their discussion several minutes later into the class period than had the whole class; consequently, the whole class had 7 minutes remaining in which to discuss the problem while the small group had only 2 ½ minutes remaining. Both the whole class and the small group spent the remainder of their discussion time on the loop question; neither reaching the free-exploration portion of the activity sheet.

Incidentally, the drawback for the small group students did not appear to be a lack of teacher support. Teacher B appeared to keep close track of the progress of each of the groups. She stopped by the small group on camera five times and was observed passing by and looking at their work at other times when she did not stop. But, it should be noted, the teacher moderated the whole class discussion *in response to the amount of time she had left*. After the discussion about the loop had continued for 5 minutes, the teacher shifted gears and spent the last 2 minutes delivering a mini-lecture to wrap up questions about this problem. In that class, the students appeared to be satisfied by the end. In comparison, the small group appeared to be close to an understanding of the loop problem when they had to stop but did not have time for a clear resolution. Although it is unknown what happened in the other small groups, the performance of the class as a whole on the related marble problem on the post-test would indicate that few if any of the small group students gained an understanding of the initial conditions required to get the skater (or marble) around the loop.

After these classes, the teacher modified the activity sheet to make the instructions concerning the key features more prominent. The next case study comparison looks at what happened in her advanced placement classes the following year, when they used the revised activity sheet.

d. Year Two: Advanced Placement Physics (Teacher B)

These case studies are, like the previous two studies, of advanced placement classes conducted by Teacher B. As during Year One, Teacher B taught the Gravitational Potential Energy lesson sequence as a one-period lesson, but this year the length of time seemed about right for the AP classes. The activity sheet and lesson plan had been tweaked in response to difficulties the teacher had observed during the first year. Though the activity sheet remained similar to the Year 1 version in Appendix B, during the Year 2 lesson, the pie chart was de-emphasized and much more attention was focused on the animated energy bar graph. Features identified as key were the GPE reference line and the animated energy bar graph. The research questions will be discussed in the order 3, 2, 4, 5,

i. Whole Class Condition: Year Two AP Teacher B

During the whole class discussion, the teacher used the simulation to answer student questions at points where she had using equations for that purpose the first year. The class appeared interested in the simulation and ready for the concepts presented. There was a fair amount of student-student discussion in low tones, some of which concerned prediction and probing questions asked by the teacher. Even though the teacher operated the simulation, which was projected onto a Smart Board in front of the class, she showed the students how to navigate to it so that they could operate it later on their own. Throughout the lesson, she offered many suggestions for manipulating the simulation. The discussion turned to topics beyond those on the activity sheet, such as how thermal energy could be produced in a frictionless environment. For instance, at one point the teacher repeatedly dropped the skater in the simulation onto the ground (to the laughter of the students) so that students could look carefully at the effect on the thermal energy bar in the animated energy bar graph.

On five occasions, the teacher explicitly invited students to turn to their neighbors to discuss their answers or predictions as they wrote on their activity sheets. The length of these discussions ranged from 7 seconds to 2 ½ minutes.

Research Question #3: Response to conceptual difficulties and misconceptions. There was no evidence for any strong student conceptual difficulty during this lesson, though S2 expressed some hesitation over the idea that thermal energy could be pertinent to their discussion of total energy.

142 T: Why won't total energy get less if I'm losing kinetic and potential?
143 S1: 'Cause you gain thermal.
144 T: Because I gain- So total does include not just kinetic and potential?
145 S2: Does it?
146 T: Does it?
147 S2**:** I think it's not pertinent because (*inaudible*).
148 T: Well, let's see what they mean by that word 'total', shall we?

The teacher then used the animated energy bar graph to help answer the question. Although the student's explanation of his reasoning was not audible, his puzzlement appears to have been at least partly conceptual, and the teacher's response was coded *response to conceptual difficulty*. Total time spent on this episode (student expression of puzzlement plus response) was about 45 seconds.

Three times the teacher appeared to be addressing a misconception where the students had shown no evidence for being aware of any difficulty. At one point, a student asked whether the skater's behavior on the track with the loop was like that of a marble on a marble track in that the shape of the track doesn't matter, "it gets to the same height it started at and then it goes back down?" If the loop situation were exactly analogous to a hilly track situation, this would imply that the skater would always make it around the loop as long as he started at a height at least equal to the height of the loop; this is not the case. (At the point the skater begins to move below the track rather than above it, additional factors come into play; gravity is no longer pulling him toward the track. See Figure 20.) Rather than explaining the reason for this difference outright, the teacher gave a hint to consider centripetal forces, then prompted the student to focus on a particular point on the track and asked a prompting question about gravity. (Boldface indicates depictive gestures.)

187 S: Is it kind of like when you are rolling a marble down the ramp and it doesn't matter what shape it is, it gets to the same height that it started at and then it goes back down?

188 T: Uh, it's probably related to that, but think about centripetal force ***[inscribes an invisible circle in the air with her forefinger]*** to go in this circle. Right? (pause) What happens if you're not going very fast at the top? ***[Points to the top of the invisible circle.]*** What's gravity doing to you?

189 S: Pulling you down.

Shortly afterward, the teacher asked for suggestions to get the skater over the loop. One student suggested a “jet pack.” The teacher chuckled, then asked what the jet pack would do. Students answered, “Increase the speed,” “Increase the velocity.” This teacher was aware of the common student misconception that a force is needed to maintain a constant velocity. She hesitated, then addressed this misconception by asking, “Would you need to keep the jet pack going the whole time?” When several students correctly responded no (because the skater only needs the jet pack to increase his speed, not to maintain it), the teacher asked why not.

A student response to this question led to the third episode. The student’s answer, “Be greater than the potential energy,” again indicated the presence of the misconception that the skater should always make it around the loop as long as his initial height was at least as high as the top of the loop. The teacher addressed this misconception:

202 T: Is it enough to be greater than the potential energy at the top? In other words, is it enough just to have *some* kinetic energy left when you get to the top?

203 Ss: (*inaudible*)

204 T: There is some amount of- (*pause*) speed we need, right?

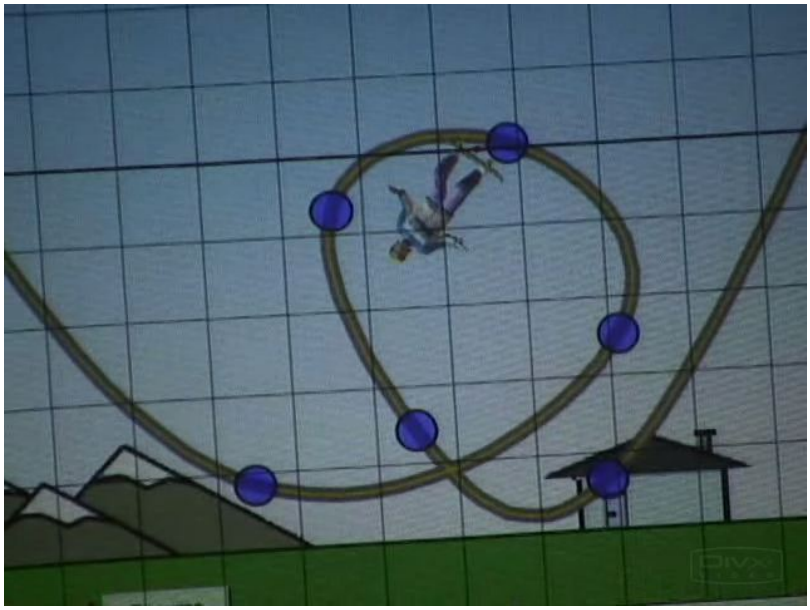


Figure 20: Loop problem.
T: "What happens if you're not going very fast at the top?"

This time, rather than mentioning centripetal force explicitly, she asked a prompting question about PE and KE to help students realize that the skater needed to have enough speed left when he reached the top of the loop that his centripetal acceleration would be at least as large as the acceleration of gravity.

These episodes each lasted between 15 and 30 seconds. Together with the episode of puzzlement, less than 2 minutes were spent on student difficulties during this 42 minute advanced placement class discussion, or about 4% of the discussion time.

Research Question #2: Discussion about key concepts: Energy with zero or negative values. Students did not appear to have a problem with the idea that TE or PE could equal zero or be less than zero. Discussion occurred during three episodes.

The first episode occurred during discussion about Question 4, when the teacher inadvertently placed the GPE reference line slightly above the bottom of the track. She then clicked a point at the bottom of the track to obtain a read-out of the energies. This number was slightly below the reference line and so showed a small negative value for the potential energy. The teacher pointed this out but none of the students asked further questions about it. The other two episodes occurred during discussion about Questions 6 and 7, which asked about zero and negative energies. Students responded to each of the teacher's queries readily and accurately and appeared not to have a problem with the concepts of zero and negative energies.

Total time spent on the key concepts was about 2 ½ minutes out of the 42 minutes class discussion or about 6% of discussion time.

Research Question #4: Support for key visual features: GPE reference line and energy bar graph. There were a number of teacher support episodes and one student support episode involving support for using and/or interpreting the meaning of the GPE reference line and the animated Energy Bar Graph.

The teacher support episodes took a variety of forms. She *gestured in the air and over the display to indicate relationships involving key features.* She also *selectively pointed out aspects of key features* using such strategies as pausing the simulation, repeatedly circling some aspect of a visual feature with the mouse, or repeating an action several times in a row and encouraging students to look carefully at transient effects. She *selectively pointed out relationships,* pointing out specific aspects of motion to look for or encouraging students to watch maxima and minima to see whether they changed. She also called attention to relationships by carefully timing transitions between scenarios in the simulation to produce maximal effect in the bar graphs. For instance, when she changed from the Jupiter scenario to the Earth scenario, she left the simulation running and made the switch just at the point the skater had converted all his potential energy to kinetic energy. The result was that the skater arrived on Earth with a large amount of kinetic energy, gained from Jupiter's large gravitational field. The bars on the bar graph went off the top of the chart and the skater sailed off the end of the track and disappeared. (See Figure 21 below.)

In the video of this episode, the teacher can be observed using at least four visual support moves in quick succession. First she *asks a prompting question* involving the bar graph, asking students to predict what will happen when she switches to the Moon scenario. She then spontaneously *suggests a manipulation*, to go by the Earth before going to the Moon, presumably so that students could compare the graph readings for Jupiter with those of a familiar environment before engaging in a more extreme comparison involving the Moon. Almost simultaneously, she uses two moves to *selectively point out relationships* that involve transient changes in the bar graph; she carefully times the switch (she had practiced this), and then gives a strong vocal reaction to the changes. At the end she asks two additional prompting questions, but those were not coded as visual support because the focus was more on interpreting the physics than on interpreting the bar graph. Again, this teacher has moved from visual support to more conceptual support. Although there were at least four visual support moves used, this 28-second sequence was counted as a single visual support episode.

The student support episode occurred when a student noticed a feature that had been added to the simulation since the last time the teacher had used it. When the teacher expressed frustration that she could not slow down the simulation to see the changes in the bar graph more clearly, a student responded, “Actually, you can.” The class laughed and the teacher immediately began to use the new slow motion feature.

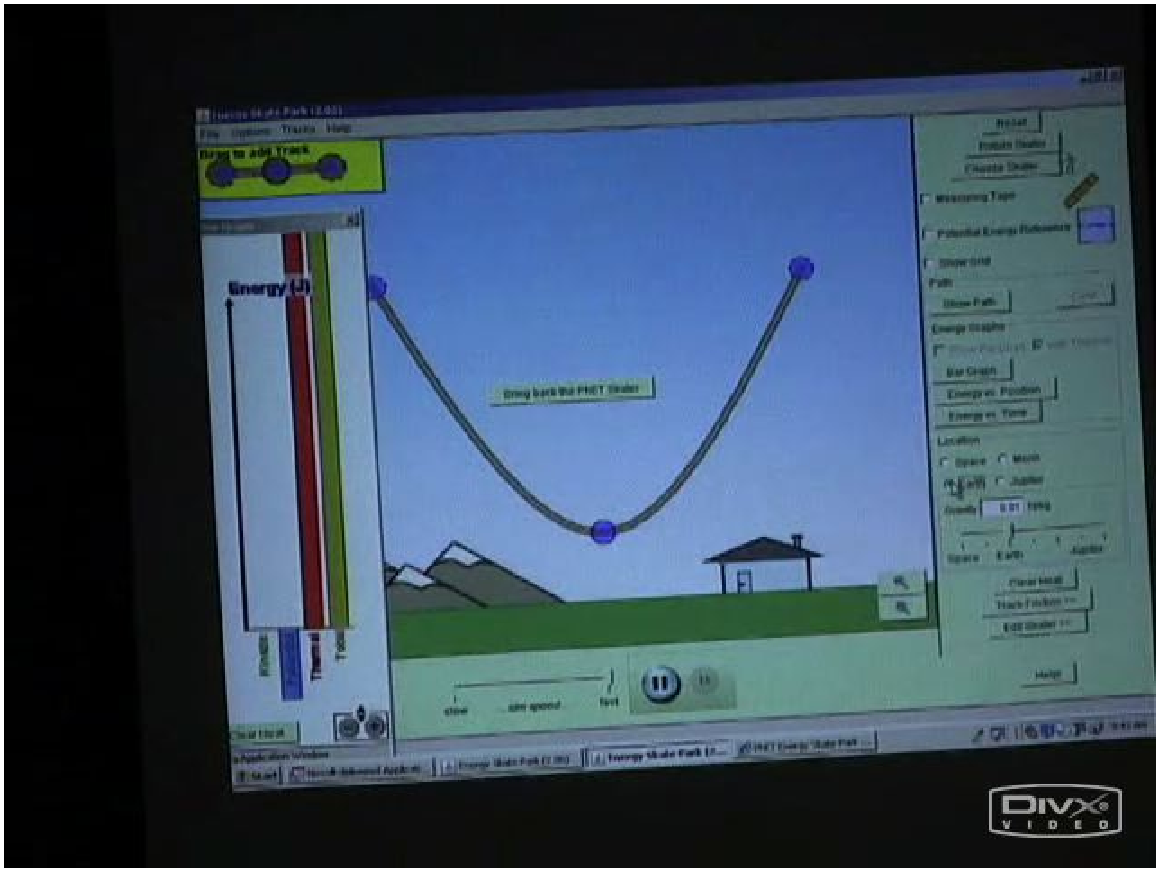


Figure 21: Jupiter Episode (Teacher provides visual support)
“Ooooh, what happened?!” Skater has disappeared, potential energy is negative, and thermal energy has gone off the chart.

There were 18 teacher episodes and 1 student episode involving visual support for using or interpreting the key features, for an average of about 27 episodes an hour.

Research Question #5: Recognition and/or use of key visual features: GPE reference line and energy bar graph. The teacher made almost continual use of the key features during this discussion, referring directly to them or to the information being indicated by them 32 times. She frequently used the cursor to refer to some specific aspect of a feature. The students referred to the key features 5 times. Four of those were in response to teacher questions that were fairly open ended and did not specifically ask them to refer to the key feature, such as, “How would you… ?” or “Why do you think…

?" The fifth time, a student mentioned a key feature unprompted by any question from the teacher.   The key features were used much more frequently than in the Year 1 AP whole class discussion (53 episodes per hour compared with 31 per hour the previous year). Part of this may reflect a change in the focus of the lesson, but may also reflect a change in the teacher's understanding of what kinds of visual support were needed by even the advanced placement students. Follow-up discussions with the teacher support the latter interpretation.

ii. Small Group Condition: Year Two AP Teacher B

As with the matched class conducted in the whole class setting, the 1-day lesson sequence seemed about the right length of time for the small groups in this AP class. The small group on camera appeared to be unusually well functioning among the small groups observed in this study. The three students in the group supported each other to come to understandings, sticking with each question until the questioner was satisfied. At one point, although two of the students were satisfied with an answer, one of them gamely manipulated the simulation for the third student to support him in continuing to explore the question. Neither of this student's partners tried to interrupt his process at any time, but appeared content to let him reach his own sense of closure on the question.

There were a number of instances of what appeared to be mutual support, in which surprise, reasoning and exploration were distributed among the three group members to the extent that it was difficult to distinguish between *difficulty* and *response*. Rather than episodes of pronounced frustration, more often observed were instances where a student said something such as, "Wait, I wanna know- " and grabbed the mouse, willing and able to use the tools at hand to investigate the issue. Some of these manipulations produced unexpected results that the students then discussed until they felt they understood them.

The teacher stopped by the small group 6 times during the half hour discussion and was present for a total of 2 ½ minutes. The students on camera finished their activity sheet several minutes before the end of the small group discussion time and explored the simulation freely until the teacher called the class back to their seats.

Research Question #3: Response to conceptual difficulties and misconceptions. There appeared to be some conceptual difficulty during four episodes; two of these episodes will be described. The second transcript excerpt is longer than most, but it will be referred to multiple times to illustrate points in connection with the other research questions below.

When reasoning about the effect of moving the GPE reference line, S2 appeared uncertain. Although he did not disagree with the answer suggested by his group partners, he asked to see for himself what would happen if the GPE line were raised above "where the skater goes." In response to S2's request, S1 moved the line about half way up the skater's path and the students watched the changes in the bar graph. But when S3 suggested that the line be put up at the top of the half pipe, completely above the skater's path, S2 objected, saying, "Well, then he won't move." S2 appeared to have the misconception that if the skater had no positive potential energy, he would not move.

Before responding to S2's misconception, S1 and S3 expressed some surprise at the behavior of the bar graph.

187 S1: It *does* go below. *(referring to energy bars going negative)*

192 S3: Because there is negative energy, so. *(pause)* That's cool.

At this point, S2 joined S3 in his earlier suggestion to move the line all the way to the top of the half pipe and S1 did so. As the three students watched the result on the bar graph, this time only S2 exhibited surprise. (Boldface indicates a depictive gesture.)

196 S2: That just affects the graph, doesn't affect what he's doing.
197 S1: Yeah. It wouldn't.
198 S2: Right.
199 S1: It's really *(inaudible)*.
200 S2: Potential energy reference- it's only (*inaudible*).
201 S1: No. Potential energy, you just call relative to the next- ***[hand held horizontally, moves it up as though moving to a higher level]***

S3 agreed with S1's comment and they moved on to the next question.

Although all three students seemed surprised initially, from Line 197 onward S1 appeared to be responding to S2's misconception. The episode lasted about a minute, although, as described above, part of the intervening discourse had more of the flavor of mutual surprise and support.

The second, and longer, episode to be discussed had both conceptual and perceptual aspects. The students were trying to interpret what they were observing in the animated energy bar graph as the simulation was being run with friction turned on. Specifically, they were trying to interpret the relationships between the changes in the different energy bars. S2 argued that a faster velocity would mean more friction, and therefore the skater's thermal energy would be *proportional* to his KE. S1 and S3 argued that the skater would lose KE due to friction, and therefore the thermal energy would be *inversely proportional* to KE. They then began to discuss whether it might be the *rates of changes* of the energies that were proportional to each other rather than the *amounts* of those energies.

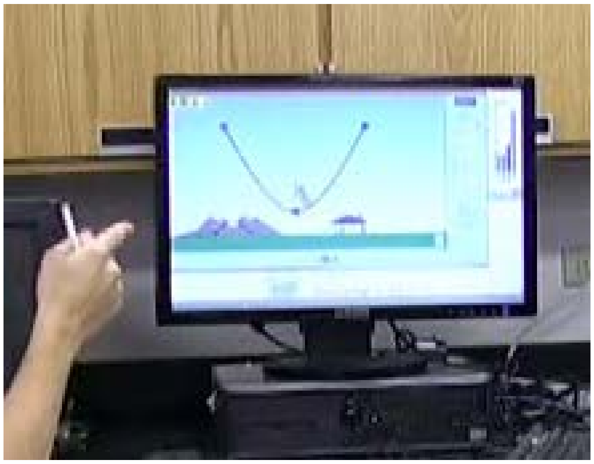


Figure 22: Friction Episode
Students try to interpret changes taking place in the bar graph at upper right of the screen. “If you look at how it increases… ”

When the simulation is in the configuration being explored by the students, many different kinds of changes occur in the energy bar graph at once, with each energy bar behaving differently. The relationships between these changes are difficult to see even when the simulation is played repeatedly as these students did. Perhaps more visually salient than increases and decreases in the heights of the energy bars, along with a general decrease over time for the maximum heights of two of those bars, is a general visual impression that the energy bars are, in some sense, ‘changing direction’ as they move. In this sense, the thermal energy bar moves upward the *fastest* at the point the KE bar is *changing direction* from up to down. At that point the PE bar is doing the reverse, changing from down to up. However, with this visual interpretation, one must observe and interpret the timing of *second order changes* in two visual elements relative to the *rate of change* of a third visual element. Perhaps it is not surprising that, though many elements of their reasoning were correct, the students had difficulty using the visual representation to definitively rule their ideas in or out.

This episode lasted less than a minute.

361 S3: I need to turn on the friction.
362 S1: Hold on. *(Looks at his activity sheet.)*
363 S2: Thermal *(inaudible)* goes up. Proportional to his kinetic energy.
364 *(Pause as they watch the simulation. Because skater has a lot of friction, he comes to a stop in a few seconds; bars on bar graph stop moving, show a final reading of no KE and a lot of thermal energy.)*
365 S1: Proportional to the decrease in kinetic energy.
366 S3: Because he is just losing so much.
367 S2: Well, but it's also proportional to the kinetic energy itself because the faster he is going, the more friction. If you look at *[points, apparently toward bar graph]* how it increases- *(Figure 22)*
368 S3: Well, the rate of change is proportional to the-
369 S1: The rate of change is proportional, but the amount of thermal energy isn't. It's, I guess it's *(clears heat and replays simulation)* inversely proportional.
370 S2: The, yeah, well, no, the *[pointing to display]* derivative of it is proportional.
371 *(Skater comes to rest again.)*
372 S3: It's kind of actually-
373 S1: Pretty cool.
374 S3: Maybe we should have a little less friction.
375 *(S3 decreases the friction and plays the simulation. The students watch it play.)*
376 S2: Seems like he's going faster- *(pointing to the simulation)*
377 S3: Yeah, obviously.
378 S1: It increases more. Yeah. But it's not-
379 S2: Yeah.
380 S1: -it's not kinetic, its thermal is changing. Anyways.
381 *(They move on to the next question.)*

In this episode, perceptual difficulties appeared to contribute to conceptual challenges. The students engaged in interesting reasoning and employed the visual affordances of the simulation to investigate their question. Although they all appeared satisfied with their perceptions by the end of the episode, coming to an agreement about their observation pattern, it was not entirely clear that the conceptual issue had been resolved for all three of them.

Response to misconceptions and/or conceptual difficulties in this small group discussion totaled about 3 minutes for the four episodes, or 11% of their discussion about the activity sheet questions.

Research Question #2: Discussion about key concepts: Energy with zero or negative values. The ideas that some energy values could be negative and that total energy could be zero occurred during two different portions of the discussion. The first portion was during the first episode of conceptual difficulty discussed above, when S1 and S3 mentioned negative energy (Lines 187 and 192).

The second portion of discussion about these concepts was in response to two questions on the activity sheet that explicitly asked about negative energy and zero total energy. During discussion about Question 6, which asked about the possibility of negative energy values, S3 indicated that there could be negative *kinetic* energy as well as negative potential energy. S1 pointed out that, according to the equation for KE ($KE = \frac{1}{2} mv^2$), this would imply that there would have to be a negative mass, and therefore KE could never be negative. However, S3 wanted to move the GPE reference line to the top of the half pipe and investigate. His partners gamely did this and they looked at the animated bar graph, at which point they all agreed that it is possible for the PE to go negative but not possible for the KE to do so.

Immediately after this, in response to Question 7, the three students agreed that the total energy could equal zero if the PE were negative enough, even if KE were present. Several seconds later, just as his partners indicated that they were ready to turn to the next question, S2 suddenly realized that it was not only PE that could be negative; total energy, too, could be negative if PE were negative enough.

Total discussion about these concepts was only a little over a minute, or 4% of the discussion time, but the students seemed to be satisfied with their answers before they turned to other topics.

Research Question #4: Support for key visual features: GPE reference line and energy bar graph. It was a challenge to determine a specific number of episodes of support; the effort prompted several cycles of refining the coding criteria that had been developed up to that point. (The refined codes were then applied to all of the Gravitational Potential Energy transcripts.) After these cycles of refinement, 8 student episodes and 2 teacher episodes were coded as supportive of key visual features.

One of the teacher episodes occurred when the students had the bar graph up and S2 expressed uncertainty about what it was that he was seeing. The teacher, stopping by their group, reminded them that they could pause the simulation in order to see what was going on, thereby *suggesting a manipulation to help them make use of the feature*.

In the second teacher episode, the students were puzzled because they could not see any energy represented on the graph. They realized that this was because the amount of energy in the system was too low for the bars to register on the screen. They were about to change their skater-track configuration to increase the total energy when the teacher stopped by. Again, she *suggested a manipulation to help them make use of the feature* by pointing out a previously unnoticed zoom feature on the animated bar graph. This feature was immediately put to use by the students, enabling them to examine the changing distribution of the very small amount of energy in their current system.

The student support episodes occurred occasionally throughout the small group discussion and consisted primarily of suggestions for manipulating the simulation. Some of these suggestions appeared to be one student attempting to help another student reach some specific understanding, but several of the episodes could probably best be described as “mutual support episodes,” with students varying in their levels of understanding but none clearly taking on the role of helper. However, the goal in each episode appeared to be to offer each other support in making better use of, or in better interpreting aspects of, either the GPE reference line or the energy bar graph. An example of mutual support was the long episode of student difficulty discussed above, when the students were trying to figure out how the different bars in the bar graph were changing relative to each other. when friction was turned on. In those twenty lines of transcript, only a single visual support episode was coded:

367 S2: Well, but it's also proportional to the kinetic energy itself because the faster he is going, the more friction. <u>If you look at *[points, apparently toward bar graph]* how it increases-</u>

In the underlined utterance and gesture, S2 has *selectively pointed out* a specific change in the bar graph to encourage the interpretation of its meaning. However, the entire episode from Line 361 through Line 380 could be regarded as involving mutual visual (and conceptual) support between S1, S2, and S3.

Eight student and two teacher episodes were coded as supportive of key visual features, for an average of 19 support episodes per hour.

Research Question #5: Recognition and/or use of key visual features: GPE reference line and energy bar graph. The students in this group made extensive use of the key features in the simulation, often when not expressly instructed to. When conceptual questions arose, they frequently turned to the simulation and used these features to investigate and reach a new degree of understanding. This can be seen in both episodes of conceptual difficulty discussed above, Lines 187-201 and Lines 361-380.

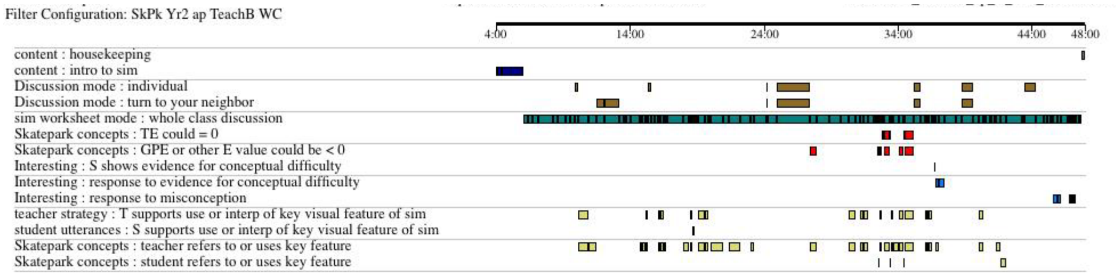
The teacher referred to the energy bar graph once when stopping by the small group, when she pointed out the zoom feature. (She had also briefly pointed out the key features when introducing the simulation before discussion began. Those episodes were not counted.) The students referred to or used the key features 30 times. This never appeared to be in direct response to a teacher suggestion and did not seem to correlate positively or negatively with the times she stopped by the small group. This lack of correlation with the teacher's presence was very unusual among the small groups observed for this study.

Other comments. The teacher had changed the activity sheet after Year One in response to student difficulties. It would have been of interest to discuss the Year One and Year Two small group interactions with these changes in mind. However, the small group on camera the first year had some unusual technical difficulties and the small group on camera the second year was probably the highest functioning in the entire study. These differences made any effect of changes in the activity sheet difficult to speculate upon with respect to the small groups on camera.

iii. Comparison: Year Two AP Teacher B

In the videotape code maps (Figure 23), the transcript segments run chronologically from left to right, spanning the time when the students were working with the animations and animation activity sheets. Color blocks below each transcript segment denote the codes assigned to that segment; the codes are listed on the left.

Whole Class



Small Group

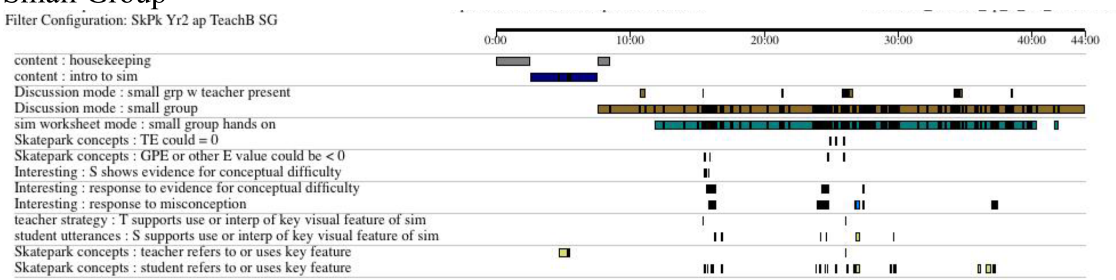


Figure 23: Videotape code maps: Year Two AP Teacher B
(Each timeline spans 44 minutes of videotape, not all of which was taken up by classroom discussion)

Table 41: Videotape coding results: Year Two AP Teacher B

	Whole Class Format	**Small Group Format**
Time provided for activity sheets (Hand out until pick up)	48 min	47 min
Time provided for simulations (including intro)	45 min	41 min
Time utilized by students on camera for activity sheet questions (Starting at Q1)	42 min	37 min
Technical & other difficulties	1 min	1 min
Length of taped discussion analyzed below	41 min 42 sec	28 min 57 sec*
Research Q #2: Discussion about key concepts	Total length: **2 min 35 sec** Percentage of discussion: **6%**	Total length: **1 min 10 sec** Percentage of discussion: **4%**

Research Q #3: Response to conceptual difficulties and misconceptions	Episodes of difficulty: 1 Response length: 40 sec Response to misc w no prior evidence of diff: 3 Length: 1 min 8 sec Total: **1 min 48 sec** Percentage of discussion: **4%**	Episodes of difficulty: 3 Response length: 1 min 33 s Response to misc w no prior evidence of diff: 1 Length: 1 min 34 sec Total: **3 min 7 sec** Percentage of discussion: **11%**
Research Q #4: Support for key visual features	Total support episodes: **19** Teacher: 18 Student: 1 Avg: **27 per hour**	Total support episodes: **10** Teacher: 2 Student: 8 Avg: **21 per hour**
Research Q #5: Recognition and/or use of key visual features	Total: **37** Teacher: 32 Student: 5 Avg: **53 per hour**	Total: **34** Teacher: 1 Student: 33 Avg: **70 per hour**

*The small group on camera took only 29 minutes to finish the questions on their activity sheet.

Discussion. As compared to the whole class condition, the small group on camera had:

- slightly smaller percentage of time spent on key concepts;
- more episodes of conceptual difficulty and **more time spent in response to difficulties,** though total amount in either condition was small;
- *lesser* frequency of visual support episodes (over a shorter amount of time; only about half the number of episodes);
- *greater* frequency of referring to the key features (over a shorter amount of time; numbers of references were actually fewer).

Notable in this comparison was that *the small group actually spent longer, and a greater percentage of their discussion time, on conceptual difficulties* than did the whole class discussion. This was very unusual among the comparisons in this study. These students also referred to key features at a greater rate than in the whole class discussion. Their rate of visual support episodes, however, was less. This may have been because the three students in the group, closely matched in ability, tended to engage in joint investigation of the visual features rather than having one student attempt to support the others. That said, their rate of visual support was still quite a bit higher than in most of the small group discussions observed.

Although the students in the small group on camera referred to the key features at a greater rate than the whole class students, this does not mean that the small group class as a whole used the features at a greater rate. Another estimate of student use of the key features can be obtained from their answers to Question #7 on the activity sheet. This second year, the teacher had split this question into two related questions, which were frequently answered as a unit by the students. (That is, their answers to Question #7 frequently referred to Question #6 or the two questions were bracketed and a single answer given.) The two questions asked whether PE or TE could ever be less than zero. Most of the students in both classes gave evidence in their answers for having used both the GPE reference line and the energy bar graph in their reasoning, ranging from 80% to 100% of the students. The only marked difference between the two classes was in evidence for use of the *relationship* between the two features; 95% of the students in the whole class condition versus 48% of the small group students gave such evidence in their activity sheet answers.

Among all the discussions observed as part of the project, these two could be considered exemplary. The small group on camera functioned very well and used the simulation to investigate their questions about the physics content. The whole class also functioned well, appearing alert and interested, responding readily to teacher questions while occasionally raising their own. Student questions occasionally moved the whole class discussion beyond the topics on the activity sheet. For instance, one student asked whether the thermal energy that appeared on the bar graph when the skater was dropped onto a frictionless track was due to work done by the normal force. This was an impressive question that the teacher paused briefly to consider before turning back to the activity sheet.

The low rate of discussion about the key concepts in both of these discussions may reflect the fact that the students in these AP classes appeared to have little trouble with these concepts. This may also explain the small amount of time spent dealing with conceptual difficulties.

Although time on task available to the two classes was within five minutes of each other, the small group on camera finished several minutes before the end of the time allowed for them, resulting in 13 minutes less time on task for that particular group than for the whole class students. However, because these students appeared to come to an understanding of the concepts quickly, they may not have needed any more time.

Pre-post gains were not significantly different between the two groups: $t(33) = 0.51$, $p = 0.62$, $d = 0.16$. Interestingly, the teacher's perception was that the whole class had not gone as well as the small group class. During a follow-up interview she stated, "The activity sheet is so tailored for Small Group that it doesn't work as well for Large Group and slows us down. ... Large Group was more problematic than Small Group. We did some major digressions at the beginning, very valuable, but then we never got to the 'imagine' thing, didn't give them enough time to talk about the loop thing." The teacher was referring to the last two items on the activity sheet, one of which asked the students to consider a skater going around a loop in the track. However, on the related explanation problem on the pre-post test, which asked whether a marble would make it over a loop and why, the whole class students actually had greater gains than the small group students, 43% to 26% respectively. This was not the only instance in this study where a teacher stated that a small group class had functioned better when the pre-post results did not reflect this.

Also in the interview, the teacher mentioned that the simple act of dropping the skater onto the ground and watching the energy bar graph react had appeared to be a useful activity for the whole class discussion, and she could imagine adding that to activity sheets for small groups in the future. This is one of several instances observed during this study where a teacher picked up a useful idea for future *small group activities* from interactions with students in the *whole class format* and vice versa.

C. Videotape Analysis: Projectile Motion Lesson Sequence

1. The Projectile Motion Lesson

The projectile motion lessons were originally planned to center on the *Projectile Motion Simulation* from Virginia Tech (Figure 24), available at *http://galileoandeinstein.physics.virginia.edu/*). This simulation was selected ahead of time by the teachers from freely available on-line sources and was used to target the understanding of the relationship between the angle and range of a projectile. However, teachers also wanted to address the independence of vertical and horizontal components of motion, and on-line simulations to address this concept in the way the teachers wished appeared to be lacking. Therefore, I used Pacific Tech's Graphing Calculator to design simple animations to supplement the Virginia Tech Simulation. Three of these animations were saved as QuickTime movies and uploaded to the school servers for use within the lesson sequence (Video Clips 1-3). The teachers tended to refer to the simulation as the “Galileo Simulation” and the three animations as the “Projectile Animations.”

The lesson sequence lasted between one and three class periods depending on the level of physics and the teacher. Although materials varied slightly for each level of physics, within each matched set, the teacher used the identical simulation and animations (Figure 24, Video Clips 1-3), prediction sheets and activity sheets (Appendices D-F) and other materials. Prediction sheets were used immediately after the pre-test but before the lesson in order to invite students to imagine and predict several effects of projectile motion. (These sheets were not analyzed for the project; they yielded much the same written data as the pre-tests, though they served a different purpose for the students.) Each teacher tossed small balls around the room as part of an introduction to the topic.

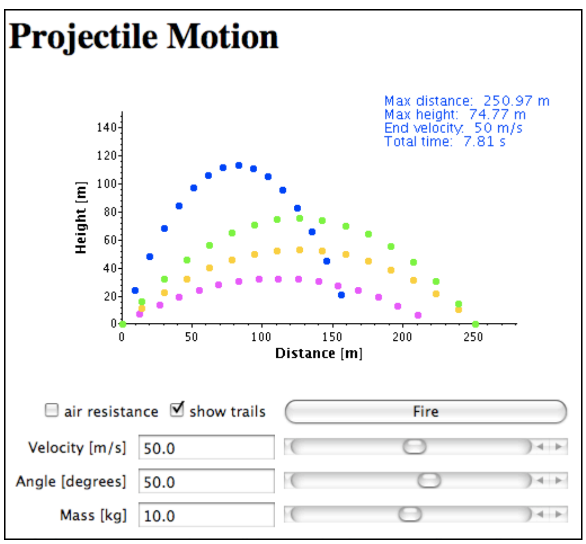
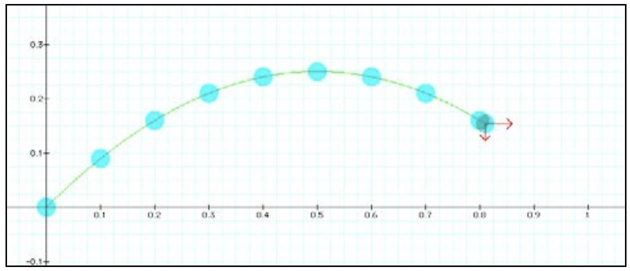


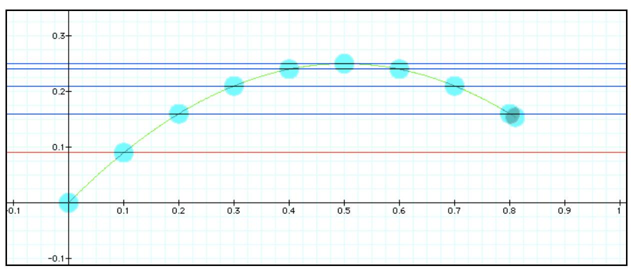
Figure 24: Projectile Motion applet, the “Galileo Simulation”
http://galileoandeinstein.physics.virginia.edu/more_stuff/Applets/ProjectileMotion/jarapplet.html

Used in Part I of the lesson sequence, this simulation creates a motion map of different projectile trajectories. Variables can be changed through data entry fields.



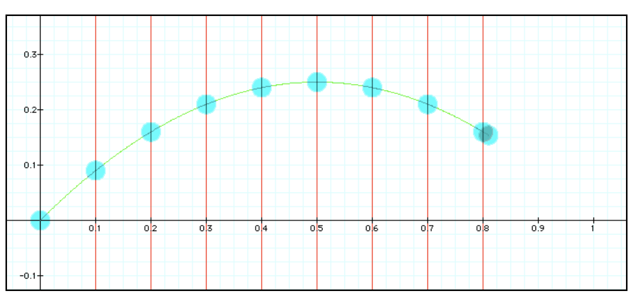
Video Clip 1: Vectors Animation (supplemental file)

In the Vectors animation, used in Part II of the lesson sequence, dots appear on the screen sequentially at equal intervals of about one second, creating a motion map. Animated vector components in red represent *x* and *y* components of velocity. QuickTime controls allow playing at various speeds, stepping through the frames individually, manually moving the projectile by dragging a slider forward or backward, pausing and looping.



Video Clip 2: Lines Animation I (supplemental file)

Lines Animation I was used along with the Vectors Animation. As dots create a motion map, lines appear at equal time intervals to show the progress of the projectile in the vertical direction. The controls are identical to those in the Vectors Animation.



Video Clip 3: Lines Animation II (supplemental file)

Lines Animation II was used along with Lines Animation I and the Vectors Animation. Lines appear at equal time intervals to show the progress of the projectile in the horizontal direction. The controls are identical to those in the Vectors Animation.

As described in Chapter III above, an exploratory study conducted during a previous year using the same simulation and animations had identified several stumbling blocks for the students. Classroom videotapes and follow-up interviews with students from the exploratory study revealed that an issue for many of them had been their difficulty identifying a concrete cause for the acceleration in the vertical direction. However, an even greater issue for some students had been their difficulty identifying a concrete reason for the *constant velocity* in the horizontal direction. (Identifying the *lack of a cause*, i.e., lack of a force, as the reason for constant velocity had proved to be problematic for many of these students.) Also, correctly identifying vertical and horizontal representations in the simulation and animations and mapping them to the phenomena they represented had proved to be an unexpected problem. All of these issues had been in particular evidence during use of the three animations, although they had appeared during use of the simulation as well. Another unexpected issue had arisen in conjunction with the animations when it became clear that few students could correctly identify the changing length of the vertical velocity vector (Video Clip 1) as an indicator of acceleration.

Classroom observations during the two years of the main study (following the year in which the exploratory study was conducted) suggested that these same issues continued to be in evidence during all portions of the lesson while remaining more clearly evident during the animations portion. This suggested focusing on the animations portion of the lesson sequence for deeper analysis even though the animations were very simple and had limited interactivity. There was another advantage to focusing on this portion of the lesson: in all lesson sequences analyzed below, the animations and animations activity sheet were used during the last class period of the sequence. Therefore, the length of time for this portion of the lesson was similar across all lesson sequences—unlike the length of the simulation portion of the lesson—and lasted no more than one class period in any of them. I decided to transcribe all class periods that occurred between the pre- and post-tests (between one and three periods for each class) and examine them for the presence of discussion concerning the key concepts as well as for moments of student confusion or conceptual difficulty, but to subject the last period of each lesson sequence, the Animations period, to deeper and broader analysis. Summaries of Period One (and Period Two where applicable) precede the analysis of the Animations period for each class.

As with the Gravitational Potential Energy lessons, the camera was viewed as a proxy for the experience of an individual student in the Projectile Motion Lessons. For whole class discussions, the camera took the position of a student in the back of the room. For small group discussions, the camera moved with the students as they broke into small groups and assumed the position of a student in one of the small groups. Therefore, the camera was focused on only one group at a time and taped what an individual student might have seen and heard. The coded transcripts, then, can be thought of as reflecting what an individual student might have been exposed to during the course of the lesson.

2. Research questions and Transcript Coding Criteria

For videotape analysis, I used elements of the constant comparative method to progress from noting observations throughout substantial portions of transcript data, to identifying patterns in observations, to defining codes that could be used for selective coding of transcripts. The codes developed to address Research Questions 2-4 in the context of Projectile Motion videotape data are below. For the Projectile Motion lessons, Research Question 5 was best addressed by analysis of the student activity sheets, which were designed with that question in mind. (See Chapter VI Section C.) Therefore, the Projectile case study descriptions of the videotape analyses address only Research Questions 2-4, but the comparative analysis of the whole class and small group case studies in each set will refer to the results of the student activity sheet analyses to address Question 5.

a. Research Question 2: To What Extent do Students and Teachers Engage in Discussion About Key Concepts While Working with the Projectile Motion Animations?

An issue that emerged in preliminary analysis was that of student difficulty in identifying concrete causal factors for different aspects of projectile motion. The key concept selected for analysis was *why* projectiles behave as they do, explained in concrete terms. Any attempt to discuss this issue was of interest even if the suggested causal factor was not correct.

Code: Student or teacher mentions concrete causal factor: Student or teacher asks question about or mentions a concrete explanation as to why some aspect of the phenomena in the system under discussion is occurring.

Concrete explanations can be distinguished from explanations constructed solely in terms of kinematic relationships or equations. If a student discussed the lack of a cause (*e.g.*, absence of forces) resulting in lack of an effect (*e.g.*, lack of acceleration), this was also counted. The transcripts were coded for any mention of a concrete causal factor for any characteristic of projectile motion. Concrete factors suggested in class discussions included gravity, inertia, "force of the throw" (which could reflect the presence of an alternative conception) and air resistance. Although student and teacher responses were coded separately, it was total time spent on these discussions that was compared.

b. Research Question 3: To What Extent do Teachers and Students Respond to Conceptual Difficulties and Misconceptions Exhibited During Work with the Projectile Motion Animations?

First, episodes were flagged where a student expressed frustration, confusion, or puzzlement in connection with ideas presented within the animation, the activity sheet, or the class discussion. Then, videotape segments that fit either or both of the following codes were coded as evidence for support of student conceptual difficulties.

Code: Response to conceptual difficulty: Classroom activity following an episode flagged as exhibiting "evidence for conceptual difficulty" was considered a response if it bore some relationship to the expressed difficulty.

Code: Response to misconception: Classroom activity was considered a response to a misconception if it appeared to be an attempt to address a misconception. (There need be no videotape evidence for the actual presence of a misconception, only that the responder appeared to think it was a potential issue.) Response could be from teacher or students or both.

For either code, the amount of discussion time spent on addressing the apparent student conceptual difficulty was established. No attempt was made to separate these responses into teacher and student responses; many responses were in the nature of joint discussion with overlapping comments.

c. Research Question 4: To What Extent do Teachers and Student Support the Recognition and/or Interpretation of Key Visual Features of the Projectile Motion Animations?

Because most of the visual features in the animations could not be manipulated, the emphasis in the videotape analysis was on support for *recognizing* and interpreting the features (and relationships involving them) rather than support for *using* and interpreting the features as it was for analysis of videotapes involving the use of the Energy Skate Park simulation. Visual features in the animations identified as key were: 1) the vertical and horizontal lines in the Lines Animations; 2) the animated red arrows that represented components of velocity in the Vectors Animation; 3) the spacing between the strobes; and 4) any visual relationships involving 1 through 3. Also considered key were: 5) any visual elements or relationships (spatial or temporal) in the animations that could indicate the presence or absence of acceleration.

It was considered that a transcript segment showed evidence for student or teacher support of other students' recognition and interpretation of key visual features if either or both of the following two codes applied.

Code: Student or teacher supports interpretation of a key visual feature or relationship in an animation as listed in 1-4 above.*

*Here, by "interpretation" of a key feature, I mean the interpretation of its meaning, the development of some degree of understanding, as opposed to attaining rote knowledge of the feature or the ability to recreate a visual aspect through mimicry.

Student or teacher is observed doing one or more of the following:

1) **Selectively pointing out** some aspect of the key visual feature or relationship as part of an apparent attempt to help students interpret its meaning**;**

2) **Giving a hint** to encourage interpretation of the meaning of the key visual feature or relationship;

3) **Gesturing in the air or over the display** to indicate the key visual feature or relationship as part of an apparent attempt to help students interpret its meaning;

4) **Asking a question to prompt interpretation** of the meaning of the key visual feature or relationship.

Code: Student or teacher supports identification of a key visual feature or relationship in an animation as described in 5 above.*

*Here, by "identification" of a key feature, I mean the recognition of its existence as described in 5 above; e.g., as a possible indicator for acceleration; I do not require any degree of understanding beyond that. This could be rote knowledge.

Student or teacher is observed doing one or more of the following:

1) **Selectively pointing out** some aspect of a visual feature or relationship as part of an apparent attempt to help students identify it as an indicator for the presence or absence of acceleration**;**

2) **Giving a hint** to encourage identification of a visual feature or relationship as an indicator for the presence or absence of acceleration;

3) **Gesturing in the air or over the display** to indicate a visual feature or relationship as part of an apparent attempt to help students identify it as an indicator for the presence or absence of acceleration;

4) **Asking a question to prompt identification** of a visual feature or relationship as an indicator for the presence or absence of acceleration.

Individual visual support 'moves' were identified and counted. Generally, for either code, when any one of the four actions was undertaken in an attempt to provide visual support, it was considered a single visual support 'move.' If the teacher or student simultaneously engaged in more than one of these actions, such as *selectively pointing out a key visual feature or relationship* while simultaneously *asking a question to prompt students to interpret its meaning*, this was counted as a single move. In long support episodes, a pause for response or a shift in tactics (asking a different prompting question, for example) was considered to demarcate between moves. However, if the same move was repeated several times in a row, it was counted only once.

Because most of the visual features could not be manipulated, Research Question 5, which concerns whether students used the features, was addressed exclusively by analysis of the student activity sheets. (See Chapter VI Section C.)

3. Case Study Comparisons of Videotape Coding Results: Projectile Motion

Eleven Projectile Motion classes, comprising five matched sets, met the criteria for the study as described in Chapter IV Section B. Descriptions of videotape coding results will be organized around Research Questions 2, 3, and 4. Following the case study descriptions of the classes in each matched set, diagrams of transcript codes and tables of results will be used to facilitate a qualitative comparison of the matched classes.

a. Year One: Honors Physics (Teacher A)

Teacher A taught the Projectile Motion lesson sequence as a two-period lesson; the periods were on subsequent days. The animations and animations activity sheet were used during the second period. Each class's experience during the first period will be summarized to give an idea of what kinds of student confusion had already been addressed and how much discussion about concrete causes for projectile motion had occurred before the students encountered the Animations. The Animations lesson will then be analyzed in terms of Research Questions 2, 3, and 4.

i. Whole Class Condition: Year One HP Teacher A

Period 1 (Summary). After the pre-test, the teacher spent several minutes introducing the topic of projectiles, tossing balls around the room as he spoke. He then let the students work individually on their prediction sheets for about 5 minutes. Next, he handed out the Galileo Simulation activity sheet and brought up the simulation onto the Smart Board in front of the class. The class spent a little over 20 minutes engaged in whole class discussion about the simulation and activity sheet questions.

Response to conceptual difficulties. No instances of student confusion were noted on this first day. In some classes, students expressed surprise and confusion when increasing the mass of the simulated projectile did not resulted in a change in trajectory. However in this class, when the teacher asked the students to predict what would happen when he doubled the mass of the projectile, three students correctly responded that the trajectory would remain the same. In regard to another issue that caused surprise in many of the classes, these students predicted that shooting the projectile from pairs of launch angles equidistant from 45 degrees would result in the same range and were not surprised when that proved to be the case. No specific responses to misconceptions were noted either; however, the teacher did support the class in testing some of their ideas and questions by using the simulation.

Discussion about key concepts: Identifying concrete causal factors. There was some discussion of concrete causal factors during attempts to explain aspects of projectile motion. For instance, when the teacher asked why it was that launching different masses at a constant angle and constant velocity had not resulted in different trajectories, one student suggested inertia, saying, “because, you know, when you accelerate something, it wants to stay the same speed and it takes more force to slow it down.” Although his reasoning was unclear, he seemed to have been thinking in terms of a greater mass being both harder to speed up and harder to slow down.

Other comments. Visual support was observed that was not specific to the simulation but that had the potential to help during the second period Projectile Animations lesson as well. Keeping track of the seemingly simple terms "vertical" and "horizontal" had appeared problematic for many students in the pilot study. During the first period Galileo Simulation lesson, this teacher explicitly addressed the distinction by employing hand gestures to illustrate the difference between horizontal and vertical motion through space (Figure 25). This appeared to be at least partially effective; there was an apparent lack of confusion in these classes about which components of motion were vertical and which horizontal. However, this clarity did not extend to the meaning of the *x* and *y*-axes in the Period 2 animations, as will be seen in the next section.

Much of the rest of the Period 1 discussion involved the teacher asking the students for suggestions for how to manipulate the simulation by choosing what numerical values to input.

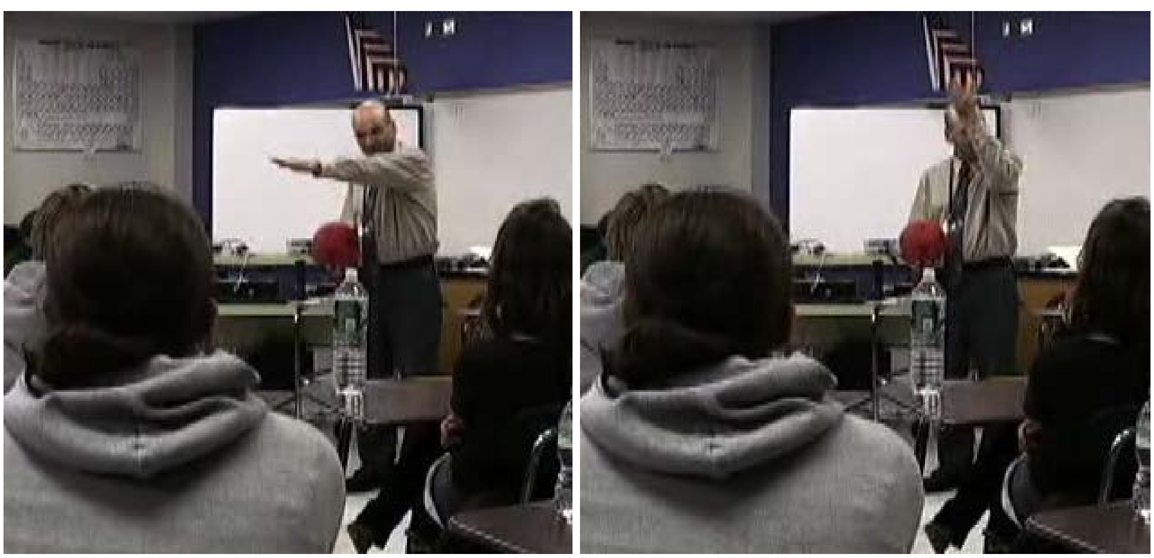


Figure 25: Teacher gestures to provide visual support
"It has a horizontal motion and also a vertical motion—together."

Period 2 (Analysis). The second period Projectile Animations lesson will be analyzed in terms of Research Questions 2-4. While the teacher was bringing up the three animations onto the Smart Board for the Projectile Animations whole class discussion, the audio was off for about 5 minutes at the beginning of the discussion and the video was off for part of that time. However, the author observed the discussion and the teacher gave a brief recapitulation when the audio came back on. According to the teacher comments and observer notes, the topic during this time was whether the acceleration represented in the Vectors Animation (Video Clip 1) was constant or varying. This continued to be the topic of discussion once the audio came back on and led to several of the episodes discussed below. The whole class discussion about the activity sheet questions continued for 21 minutes.

Research Question #2: Discussion about key concepts: Identifying concrete causal factors. Only one mention of a concrete causal factor was observed on this second day. At the beginning of the discussion, while the audio was off, a question was raised about the cause of the acceleration in the Vectors Animation. A few seconds after the audio came on, the teacher returned to this, asking, "(I)s there enough information on here that we could answer the question that S3 brings up of, is this acceleration due to gravity, or just all we can say is, it's acceleration?" Another student replied that there were no labels on the graph, with the implication that this left her unable to answer the question of what was causing the acceleration. Causal factors were not mentioned again in this discussion. Gravity had not been a central topic of discussion during the time the audio was off, so to be conservative, only the 15 seconds of discussion about gravity after the audio came on was counted as *discussion about a concrete causal factor.* This was 2% of the whole class discussion time.

Research Question #3: Response to conceptual difficulties and misconceptions. There was one point at which the teacher guided the discussion to respond to a misconception and another at which he responded to a student's apparent frustration or confusion.

Near the beginning of the discussion, a student asked whether the acceleration in the Vectors Animation was constant (correct) or varying (a common misconception). The teacher began the response by asking what constant acceleration would look like and whether there was anything in the animation that could be measured to provide an answer to the student's question. Two students suggested that the change in Arrow A could be measured. (See Video Clip 1.) The teacher then expanded on their answer, gesturing over the display and saying that if the rate of change of Arrow A were constant, then the acceleration would be constant. A student very softly said, "Constant, that makes sense." The teacher's explanation in terms of Arrow A lasted about a minute and appeared to be effective, at least for some students. This part of the discussion is discussed in more detail below. (See Lines 20-24 in the next section.)

Only one instance of mild frustration was observed. During discussion about Lines Animation I (Video Clip 2), which used lines to indicate the changing amount of vertical distance traveled during equal time periods, S4 voiced some frustration at the teacher's suggestion that the *x*-axis might not represent time. The excerpt begins with the teacher's suggestion. (Square brackets indicate gestures and boldface indicates depictive gestures.)

67 T: The one thing that I'd remind folks is that the word *strobe* suggests that those **[G]** blue circles are fixed there at equal time intervals. It doesn't tell us what that [*G*] *x*-axis is, so we can't be sure, is that a position or a time axis-

68 S4 (*interrupting*): But it says that the lines are velocity; the question has velocity in there, so that's what (*inaudible*).

69 T: *(Re-reads question and points out some of the relationships in the graph but does not directly address the student's statement.)*

70 S4: Right, I'm just saying we know it's a position-time graph because we use the term 'velocity'.

When S4 said the question had "velocity in there," she was referring to a question on the activity sheet that asked what the lines in Lines Animation I indicated about the velocity of the projectile. She argued that the phrasing of the question meant the lines *were* velocity and, therefore, the graph was a position-time graph (in which the *x*-axis would represent time). S4 was not the only student in the observed classes who was sure that the *x*-axis represented time.[9] Actually, all three of the animations were position-position motion graphs and the *x*-axis represented horizontal distance. In Lines Animation I, the lines indicated the height of the projectile (its position along the *y*-axis) at equal time intervals.

In Line 71, the teacher indicated that he now understood S4's question.

71 T: Aaah. OK. Interesting. So what *(S4)* said, is that, because the word 'velocity' appears here, that this must be **[G]** position on the *y*-axis and **[G]** time on the *x*-axis. Is it possible that there is **[G]** position in the vertical on the **[G]** *y*-axis, and **[G]** horizontal position on the *x*-axis? And that the time component is represented by **[G]** the strobes?

He then invited students to imagine what would look different in a position-time vs. a position-position representation of projectile motion. The last part of the teacher's

[9] Interestingly, Schultz, K., Clement, J., & Mokros, J., (1986); McDermott, Rosenquist, & van Zee (1986); and others have found that many students interpret graphs as pictures, apparently interpreting the temporal direction on the graph as a spatial direction. In view of this, the tendency of students in the present study to view one of the spatial directions as a temporal direction was not anticipated, but perhaps understandable given that all of their prior experiences with graphs in these physics classes had been with *x*-*t* graphs.

response will be discussed below in terms of the visual support the teacher offered. (See Line 73 in the following section.)

Total time for the whole class responses to misconceptions and difficulty was less than 2 ½ minutes, or 15% of the discussion time.

Research Question #4: Support for key visual features: Lines, arrows, spaces, and acceleration indicators. Throughout this lesson, the teacher used several strategies to support students' interpretation of key visual features in the animations. There was also a single student support episode. Three transcript excerpts involving five episodes are described. Underlined sections were coded as support for key visual features.

Concerning whether the acceleration represented in the animations was constant or variable:

20 T: Let me ask you a question: What would constant acceleration look like? If there was something you could measure on this *(display)*, what would that be? The question, by the way, there is a discussion going on about whether the acceleration is constant or not. *(Vectors animation is playing on loop.)* One person is claiming there is acceleration but it's not constant, the other one is saying no, it looks constant. If there was something we could go up there with a ruler and measure, what would it be?

In this first support episode, the teacher was *asking a question to prompt the identification of a key visual relationship as indicator for the presence of acceleration* and also *to prompt interpretation of the meaning of the relationship* in terms of whether it represented constant or varying acceleration. (He also repeated the question, but the repetition was not counted.) The relationships that indicated the constant nature of the acceleration were only implicit in the animation (Video Clip 1), and most students interviewed the previous year had not been able to identify them. However, in this class, there was a correct response to this question.

21 S5: The change in the vertical velocity component.

(Correct.)

23 S2: The change of the arrow A is the change of displacement.

(Incorrect, but a more typical answer among the students observed in this study.)

24 T: So what I hear S5 suggesting is that this arrow here *(using the cursor to indicate the v_y vector component in the animation)* is changing and if its ***[with thumb and forefinger, brackets a distance in the air of about an inch]*** rate of ***[repeats gesture]*** change is constant, that would be constant ***[repeats gesture over the display]*** acceleration. So for example, maybe we could sort of- look at it at equal intervals, here *(using cursor to indicate positions on the x-axis)*, over point one *(x=0.1)* and over point two *(x=0.2)*, or something like that, and measure how long that red line *(the v_y vector component)* is and see if the ***[with left thumb and forefinger, indicates several different vertical distances in the air]*** change is equal.

In this second support episode, the teacher made explicit which visual cues could give information about the acceleration. In the animation, if the change in length of the vertical vector component of the velocity (vertical red arrow) is the same for equal intervals along the *x*-axis (which also happens, in this case, to correspond to equal time intervals for the projectile), then the acceleration in the vertical direction is constant. The support moves included *gesturing over the display to indicate the vertical red arrow* and *selectively pointing out a visual relationship* that a student was describing but that was not necessarily easy to see. This visual support appeared designed to help address a common student misconception that when velocity changes, it means that the acceleration is varying.

The above episode did not straighten out the confusion about what quantity was represented along the *x*-axis, and may even have increased that confusion, but it did point out the nature of the information given by the changing length of the vertical red arrow.

Later, in another episode already described above, the teacher addressed this confusion directly.

67 T: OK, so we've heard a couple of interesting observations. The one thing that I'd remind folks is that the word 'strobe' suggests that those **[G]** blue circles are fixed there at equal time intervals. It doesn't tell us what that [*G*] *x*-axis is, so we can't be sure, is that a position or a time axis- (*rhetorically*)

Although, in the next line, S4's confusion about the meaning of the *x*-axis became apparent (see Conceptual Difficulties section above), the teacher has begun to deal with the issue here. He has *selectively pointed out the spacing between the strobes* to encourage students to interpret the spaces as representing equal time intervals. He hoped this could help students see that it was not necessary for the *x*-axis to represent time in this particular graph. In doing so, he has also called attention to a feature that indicated the presence of acceleration: *unequal* spaces between the strobes corresponded to *equal* time intervals.

In a fourth episode addressing the same issue a few minutes later, the teacher suggested the following.

73 T: OK, I want you guys to keep your mind open to that. (*Referring to the possibility that the time component could be represented by the strobes rather than by the x-axis.*) Ya know, the question then might be, what would look different? On a position-time graph or a position-position graph? How would they appear different when what's represented is the motion of a projectile? Interesting. (*pause*)

The answer is that the graphs could look identical, depending on the scales chosen to represent *time* and *distance* and on the velocity of the projectile. However, in the general case, the *y*-*t* graph of a projectile would look compressed or expanded along the horizontal axis as compared to a *y*-*x* graph of the same projectile. Rather than stating this directly, the teacher *asked a prompting question*, "(W)hat would look different? On a position-time graph or a position-position graph?" This was a support move for *interpreting key visual relationships* in the graphs; it could also be considered support for *mental* visual imagery. The teacher appeared to be attempting to support students in modifying their understandings of the red horizontal arrow and the horizontal relationships in the graphs by inviting them to try *mentally* to compare graphs in which the horizontal direction represented *x* vs. those in which this direction represented *t*.

There was one student episode. When the teacher asked, "What do you notice about that component?" referring to the velocity component represented by the animated horizontal arrow, a student answered, "Stays the same." Another student then responded:

30 S6: The strobes tell you that. 'Cause the distance between each strobe in regards to *x*-values are equal.

This student has *selectively pointed out* a different set of features than the ones the teacher was asking about *to support other students in interpreting a relationship* in the display. The student utterance appeared to be intended to help other students discover information derivable from the relationships between the strobes and to relate that to information discoverable from the unchanging length of the horizontal arrow.

There were 11 teacher and 1 student visual support episodes during the videotaped discussion, equivalent to 44 visual support episodes per hour. Many of the teacher episodes involved depictive gestures.

ii. Small Group Condition: Year One HP Teacher A

Period 1 (Summary). After the pre-test, the teacher tossed balls around the room as he discussed terms to be used in the Projectile Motion lesson, "range" and "hang-time." He then handed out the prediction sheets and allowed students about 5 minutes to work on them. After demonstrating the controls on the simulation and having the students break up into small groups, they had about 15 minutes remaining in the class period to work at computers on the activity sheet questions. In the small group on camera, one student acted as an authority, explaining relationships to the rest of the group as he figured them out.

Response to conceptual difficulties. The students in the observed small group expressed some puzzlement about aspects of the Galileo Animation. However, unlike in some of the other classes observed, the puzzlement during this first period of the lesson appeared to do with accurately identifying observation patterns rather than with explaining them. For instance, they asked whether the range quadrupled with increase in velocity. There were some explanations offered but these were not accompanied by any noticeable conceptual difficulty.

Discussion about key concepts: Identifying concrete causal factors. S1 suggested a concrete causal factor to help explain a pattern that had provoked conceptual difficulty in other observed classes. The pattern is that in the absence of air resistance, when mass is varied with all other variables including launch speed held constant, the trajectory does not change. S1 pointed out that it would take more energy to launch a larger mass than a smaller mass at the same speed. This concrete causal factor offered a partial, but apparently satisfying, explanation for why it was possible for the larger mass to follow the same trajectory as the smaller one.

Period 2 (Analysis). The Projectile Animations lesson will be discussed in terms of the Research Questions 2-4. For this lesson, the camera joined the same 3-person small group as it had during Period 1. S2 acted as expert, explaining the concepts to the other two students (though not always correctly), pointing to the animations and explaining what the different elements represented and why. The other two students appeared to accept S2 in the role of expert. S1 was quiet for the most part, occasionally asking for clarification before writing her answers. Although S3 seemed to accept S2 as expert, he actively reasoned aloud and appeared to understand much of what he saw.

There were 17 minutes available for the small group activity. However, the students had to get the computer and the animations up and running and they brought up the wrong animation first, with the consequence that the questions on the activity sheet appeared to make no sense. After they figured out what was wrong and found the correct animation, they had just under 12 minutes remaining in which to complete the activity.

Research Question #2: Discussion about key concepts: Identifying concrete causal factors. This small group was unusual among those observed for the Projectile motion lessons because concrete causes for the phenomena were mentioned on 5 different occasions, for a total of about 2 minutes of discussion. The first time was during the Vectors Animation (Video Clip 1), which used animated arrows to represent the vector components of velocity of a projectile. In the transcript excerpt below, the group was discussing why, as the animation plays, the horizontal Arrow B stays the same length. Boldface indicates a depictive gesture by the speaker.

59 S1: Wait, why would that be, though?
60 S2: Why, the horizontal be the same?
61 S1: Shouldn't it (*horizontal velocity component*) get smaller?
62 (*All three students look at the display.*)
63 S2: No. Because it's still traveling ***[traces invisible horizontal path over the screen]*** this way the same. It's just that the energy ***[with thumb and forefinger, indicates invisible vertical quantity]*** is being-
64 S1: -Yeah, that makes sense.
65 S2: -the energy ***[indicates vertical quantity, then moves hand over the screen]*** of going like that is being- yeah.
66 S1: The gravity is pushing down ***[gestures down]***, not this way ***[gestures horizontally]***.

67	S2: Right.

In this episode, S2, the resident expert, gave a rather inarticulate answer but S1 responded with a clear and concrete cause for the acceleration in the vertical direction and the lack of acceleration in the horizontal direction. S1's words were given added clarity by his two gestures.

At another point, S2 mentioned that "gravity is positive acceleration." This appeared to be a causal explanation for him although perhaps not for his group mates.   A few minutes later, S1 and S2 discussed the apex of the trajectory as being caused by something becoming equal to gravity, although they weren't quite clear on what that "something" was. When writing their answers on the activity sheet, twice S2 reiterated that the acceleration was downward and caused by gravity. Total time for discussion about concrete causes was slightly less than 2 minutes, about 17% of this group's discussion time (counting from when they got the correct animation up).

Research Question #3: Response to conceptual difficulties and misconceptions. There was only one point in this discussion when a student exhibited puzzlement and that was in Lines 59 and 61 in the episode in the previous section, when S1 asked why the horizontal velocity vector would not get smaller. The utterances coded as *response to conceptual difficulty* were Lines 60 and 62-67. In these lines, S2 responded by attempting an explanation in terms of energy, whereupon S1 responded to his own confusion by successfully generating an explanation in terms of the direction of gravity. Time spent was only 14 seconds, but that seems reasonable given the fact that the confusion appears to have been resolved.

There were four episodes in which one of the students appeared to be trying to respond to the misconception of another student when the other student had not exhibited any awareness of experiencing conceptual difficulty. Three of these episodes involved the question of whether there was a change in acceleration. Each of these times, S1 attempted to respond that the acceleration was not changing. The first time, he simply replied that the acceleration was not changing, but the second time, he engaged S2 in a brief discussion about it. The third time, he clearly stated, “It’s not going from up to down; it’s always going down.”

In the fourth response to a misconception, the students were discussing a activity sheet question about Lines Animation 1 (Video Clip 2), "What does the variable spacing between the red and blue lines indicate about the velocity?" These lines marked the progress of the projectile along the *y*-axis at equal time intervals. Thus, the horizontal lines gave information about the vertical component of velocity. Many students found this confusing. In this instance, the students had agreed that the lines indicated something about the upward velocity. S2 then asked a pedagogical question; that is, he appeared to ask a question for which he knew the answer in order to help S3.

151 S2: Right, so does it give you horizontal information or vertical information? (*Looks at S3. Sounds professorial, not puzzled.*)
152 S3: Horizontal.
153 S2: Well-
154 S1: It's vertical, isn't it?
155 S2: Yeah, it gives you vertical because it's saying that in ***[with fingers, brackets horizontal distance between two strobes]*** this amount of time, the ***[brackets vertical distance between two strobes]*** vertical change was this. *(Figure 26)*
156 S3: Oh OK.
157 S2: In this ***[brackets next horizontal interval]*** amount of time ***[brackets corresponding vertical distance, less than previous vertical gesture]*** the vertical change was only (*inaudible*).

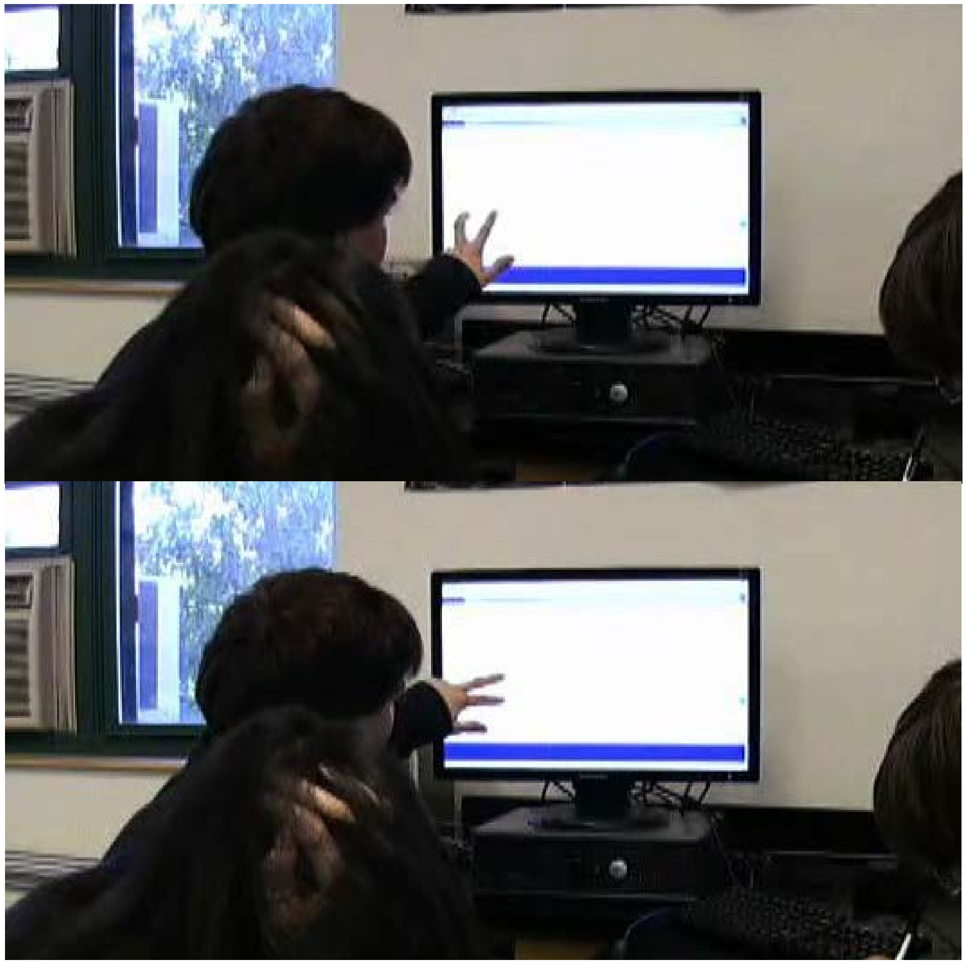


Figure 26: "In *this* amount of time, the vertical change was *this*."

S2 used gestures and words to explain the relationships on the screen to address the misconception that horizontal lines necessarily gave horizontal information. (Lines 155 and 157 were also counted together as a single episode of visual support.)

Total time spent on addressing conceptual difficulties and misconceptions was 1 min 21 seconds, or 11% of the discussion time.

Research Question #4: Support for key visual features: Lines, arrows, spaces, and acceleration indicators. There were 11 episodes when one or more students attempted to support others in interpreting visual aspects of the animation. In three of these, S2 attempted to support the other students in understanding that the change in

Arrow A indicated acceleration and that the lack of change in Arrow B indicated the absence of acceleration. Nine of these episodes were accompanied by gestures in the air or over the animation. In several of these episodes the gestures were depictive; that is, they appeared to depict an invisible shape, location, or pathway in the air. The other gestures involved the speaker pointing to some visual element in the animation.

In the following transcript excerpt, there were three visual support episodes. Square brackets indicate gestures, boldface indicates depictive gestures, and underlining indicates a visual support episode.

130 S2: *(reading)* "What does the variable spacing which is between the red lines and blue lines indicate about the velocity?"
131 S1: It's slowing down.
132 S2: Well, what's this a graph of?
133 S1: Position and time.
134 S2: Are we sure?
135 S1 & S2 *(Overlapping):* Yeah. Yeah.
136 S1: Yeah, 'cause they're getting *[holds two fingers up to the screen and brackets a small vertical distance]* closer together, yeah, the velocity is decreasing, because, if it's-
137 S2: If it- so *[points to several horizontal lines on the display]* every single one of these lines is ***[moves hand in staccato motions from left to right across the screen even though the lines are horizontal]*** a time increment. Is its ***[bent hand, fingers horizontal along one of the lines]*** height at a ***[straight vertical hand against the screen, moves it slightly to the right]*** time increment. *(Stated as a rhetorical question.)*

In Line 132, S2 *asked a prompting question* that appeared to be inviting S1 to re-evaluate his interpretation of the variable spacing of the lines, a key visual feature, although it is not entirely clear toward what interpretation S2 was trying to point.

In Line 136, S1 tried to explain to S2 why he was sure that the graph was a position-time graph. He *gestured over the display* to indicate the relationship between the lines, a key feature, apparently to help S2 discover additional information that S1 believed was derivable from this relationship. He was partly correct; in the animation, when the lines are getting closer together, this does indicate that the velocity is decreasing. This episode was also counted as supporting other students to *identify the relationship between the lines as an indicator of acceleration*. Even though S1 did not use the term "acceleration," he did say "the velocity is decreasing," and it is clear from other portions of the transcript that for him, a change in velocity did indicate the presence of acceleration. (It was not always correct to assume that the students in this study equated a change in velocity with acceleration.)

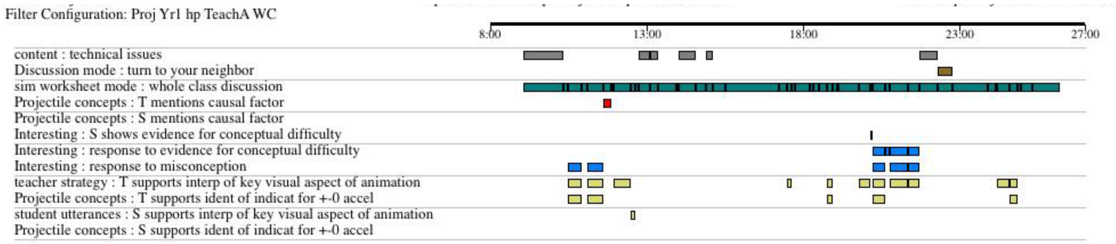
In Line 137, S2 *gestured over the display* to indicate vertical lines and strobes. This, along with his words *selectively pointing out* these features, appeared intended to help the other students discover additional information derivable from the relationship between the different heights of the projectile at different time increments.

There were 11 visual support episodes in 12 minutes of discussion, for an average of 55 episodes an hour. The teacher did not stop by this small group during their discussion.

iii. Comparison: Year One HP Teacher A

The videotape code maps (Figure 27) and table (Table 42) represent only Period 2 of the lesson sequence, during which the three Projectile Animations were used. In the code maps, the transcript segments run chronologically from left to right, spanning the time when the students were working with the animations and animation activity sheets. Color blocks below each transcript segment denote the codes assigned to that segment; the codes are listed on the left. In these videotapes, the camera was used as a proxy for an individual student; the codes can be considered to reflect what an individual student in that class might have experienced.

Whole Class



Small Group

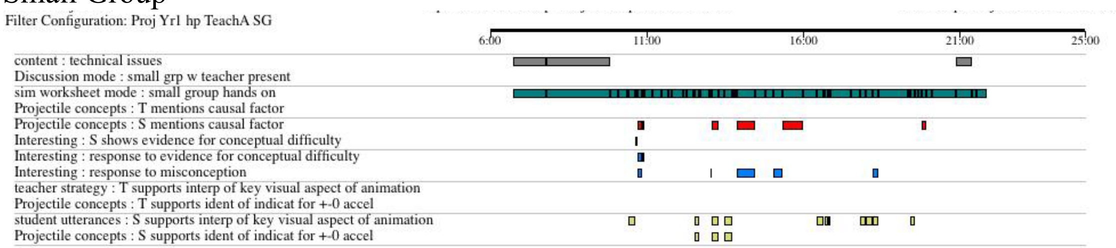


Figure 27: Videotape code maps: Year One HP Teacher A
(Each timeline spans 19 minutes of videotape, not all of which was taken up by classroom discussion)

Table 42: Videotape coding results: Year One HP Teacher A

	Whole Class Format	**Small Group Format**
Time provided for activity sheets (Hand out until pick up)	23 min	21 min
Time provided for animations (including intro)	21 min	17 min
Time utilized by students on camera for activity sheet questions (Starting at Q1)	21 min	14 min (but had wrong animation up for 2 min)
Technical & other difficulties	5 min*	5 min**
Length of taped discussion analyzed below	15 min 53 sec	12 min 1 sec
Research Q #2: Discussion about key concepts	Total length: **15 sec** Percentage of discussion: **2%**	Total length: **1 min 50 sec** Percentage of discussion: **15%**

Research Q #3: Response to conceptual difficulties and misconceptions	Episodes of difficulty: 1 Response length: 1 min 28 s Response to misc w no prior evidence of diff: 1 Response length: 57 sec Total: 2 min 25 sec Percentage of discussion: 15%	Episodes of difficulty: 1 Response length: 14 sec Response to misc w no prior evidence of diff: 5 Response length: 1 min 7 s Total: 1 min 21 sec Percentage of discussion: 11%
Research Q #4: Support for key visual features	Total support episodes: 12 Teacher: 11 Student: 1 Avg: 45 per hour	Total support episodes: 11 Teacher: 0 Student: 11 Avg: 55 per hour

* In whole class format, the first five minutes of discussion were not captured on video due to technical problems, although observer notes were recorded.
**In small group format, it took the group on camera time to get the computer going, then they had the wrong animation without realizing it. They also had technical challenges later in the period.

Discussion. This Honors Physics whole class discussion included about 21 minutes of on-task discussion time, 16 minutes of which was caught on camera. The small group observed in the matched class had 17 minutes of discussion time but the first five minutes were not productive because the group first had to get the computer going and then had the wrong animation up. The students spent time trying to figure out what was wrong rather than addressing the activity sheet questions. This appeared to be one of the risks of small group work—time can be wasted because students can make mistakes when they follow instructions. However, because I wish to be conservative about reporting any advantage for the whole class condition regarding frequency of support episodes, for the small group I counted only the 12 minutes they had the correct animations up. Because there were no episodes in the initial 5 minutes of confusion that met coding criteria, this had the effect of increasing the small group averages and percentages in the grid above.

Calculating in this way, as compared to the whole class discussion, the small group discussion had:

- **much greater percentage of discussion time about key concepts** (and greater length of time);
- more visual support episodes per minute (but about the same number of episodes);
- a smaller percentage of discussion time addressing conceptual difficulties and misconceptions.

The greatest difference observed between the two discussions was in the percentage of discussion time spent on the key concepts: discussion of concrete causal factors. However, concrete causes for the motion had been mentioned the previous day in both classes. In the whole class discussion, students had raised the issue twice and in the small group discussion, students had raised the issue once on the first day. However, it is notable that the teacher was not observed mentioning concrete causes at all on the first day, and on the second day in the whole class discussion, for only 15 seconds. Causal factors had been chosen as key concepts because they appeared to have been so important to students in the exploratory interviews conducted during the previous year. However, as can be seen here, Teacher A did not appear to put much emphasis on concrete causes for projectile motion during these Year 1 lessons. It could be that he assumed that all the students knew that gravity was the important causal factor and that it did not need to be discussed. Students in the small group on camera seemed to find it important, though, not only to remind themselves that acceleration was occurring in the downward direction because gravity pulls down, but also to establish that velocity was constant in the horizontal direction because of inertia. The importance to students of understanding the *lack of a cause* and how it can explain the *lack of an effect* (in this case, lack of horizontal acceleration) appeared to have been underestimated by this teacher during the first year of the study. The second year, he discussed concrete causal factors more explicitly, as will be seen in case studies below.

The small group also exceeded the whole class discussion in frequency of visual support episodes, though the actual number of episodes was about the same. This number of support episodes was very unusual among the small groups observed for the study. This was largely due to the single student who took on the role of "expert" and attempted to explain concepts and visual elements to his discussion partners, frequently gesturing over the screen as he pointed out visual relationships.

The teacher did not stop by this small group during the Animations Lesson discussion. Because one student took on the role of expert, it may be that this small group took on some of the characteristics more common for small group discussions with a teacher present, and perhaps even some characteristics more typical of whole class discussions. However, the student "expert" did not fully fill the role of teacher. One way this was reflected is the way in which he responded to the conceptual difficulties and misconceptions of his discussion partners. The percentage of discussion time devoted to misconceptions and conceptual difficulties in the two discussions was similar: 11% for the small group discussion compared to 15% for whole class. In the small group, most of this time was S2 addressing misconceptions *he perceived* his partners to have rather than addressing difficulties *his partners expressed*. When one of his partners did express conceptual difficulty, the "expert's" response was very brief, 14 seconds. In the whole class discussion, on the other hand, after the single episode of student conceptual difficulty, the teacher gave an initial response and then addressed the issue in two additional ways, first using hand gestures and other visual support strategies and then inviting the students to imagine two scenarios and mentally compare them. His response lasted about a minute and a half, far longer than the "expert's" in the small group.

Although the experiences of the other small groups may have differed from those of this unusual group in a number of ways, the fact that there was no significant difference between the pre-post gains in these two classes [$t(44) = -0.09$, $p = 0.93$, $d = 0.03$] suggests that there may indeed have been factors in the whole class discussion that helped counterbalance the small group affordances. Both groups had significant gains at the $p < 0.001$ level with relatively large effect sizes ($d = 1.33$ and 1.17 for WC and SG respectively).

The whole class condition had somewhat larger gains on the explanation questions of the pre-post test, 19% gains as opposed to 11% for the small group class. The concept targeted by these questions was how the velocity components of a projectile would change if the value of gravity were changed. (Answer: The vertical component of velocity would change while the horizontal component would not be affected.) These questions would appear to require reasoning about causal factors, and the greater gains for the whole class condition suggest that, even though they had had less discussion about these factors than at least one of the small groups, the whole class discussion may have provided other affordances for reasoning about such factors.

The whole class students had a lower rate of support episodes for the key visual features than the small group on camera, an unusual result in this study. Analysis of several activity sheet questions can give an estimate of the extent to which the students were actually able to make use of the features in their reasoning. These results suggest that the small group on camera may not have been as much of an outlier as they initially appeared to be; other small groups in this class also appeared to be able to make use of the features. While the whole class averaged 63% correct on their answers to the four questions, the students in the small groups averaged 75%, suggesting that both groups understood the key features and could use them in their reasoning, and that the small group students may have been slightly more successful with this.

This was the first year Teacher A had taught the lesson. As will be seen in the Honors whole class discussion he led the following year (Section c. iii, Table 44), he engaged in more visual support episodes and spent far more time on concrete causal factors than during this first year.

b. Year One: Honors Physics (Teacher C)

Teacher C taught the Projectile Motion lesson sequence to three matched classes, two using the whole class discussion format and one using the small group format. She elected to teach the lesson as a three-period sequence, using the Galileo Simulation during the first two periods in each class and the Projectile Animations during the third period. Analysis focused on the Animations lesson. The first two periods will be summarized to give an idea of what kinds of student confusion had already been addressed and how much discussion about concrete causes for projectile motion had occurred before the students encountered the Animations. The Animations lesson will then be analyzed in terms of Research Questions 2, 3, and 4.

i. Whole Class Condition #1: Year One HP Teacher C

This was the first of Teacher C's three matched classes. In this class (Class A), the simulation and animations were used in the whole class format.

<u>Period 1 (Summary).</u> After the pre-test, the teacher handed out the prediction sheets and then asked the students to pause while she went over terminology to be used in the Galileo Simulation lesson (e.g., *range:* horizontal displacement; *maximum altitude:* maximum vertical displacement; *hang time:* time between launch and landing). The students then filled out their prediction sheets, working independently. Next the teacher led an introductory discussion during which she threw balls around the room and asked students what factors were affecting the motion. They answered with several concrete causes: gravity, air resistance, applied force, and launch angle. The teacher then projected the Galileo Simulation onto a screen in front of the class and discussion turned to the activity sheet questions. They began a whole class discussion that lasted for 18 minutes, until the bell rang.

Response to conceptual difficulties. The students expressed puzzlement, bewilderment, and/or frustration at two points. The first was when students discovered that, contrary to their predictions, doubling the mass in the Galileo Simulation while holding the launch velocity constant did not result in a change to the trajectory. Several students were heard repeatedly saying, "I just don't get it." A four-minute whole class discussion ensued in which students and teacher actively participated in addressing the difficulty. One strategy the teacher used was to respond to student questions by suggesting that the class use the simulation to investigate them.

A second incident of conceptual difficulty occurred when students realized that firing the simulated projectile at two different launch angles resulted in the same range for the projectile. This time the teacher did not appear to change what she was doing in response to student comments such as, "Uh, I'm confused!" The activities the class was already engaged in were designed to address this very issue and the students soon discovered that complimentary angles produce the same range. Identifying this pattern seemed to satisfy some students, but one student wanted to know how complimentary angles produce this effect. This question was raised after the bell, however, and the teacher merely smiled and shrugged. Because she had allowed a second day of work for this simulation, she knew that the students would be engaging in further investigation concerning this issue on the following day.

The discussion spent a total of about five minutes in two episodes of addressing conceptual difficulties. Notable in both episodes was strategic decision-making by the teacher concerning when to bring closure to the conceptual issues.

Discussion about key concepts: Identifying concrete causal factors. Before the simulation was brought up or the activity sheets handed out, the teacher asked the class what factors might have an impact on the way a projectile moves. Students suggested four causal factors: gravity, air resistance, applied force, and launch angle. These factors were mentioned again during the class's use of the simulation, twice by students and once by the teacher; however, less than half a minute of the discussion was spent on this.

<u>Period 2 (Summary).</u> In the second period of the 3-period lesson sequence, the teacher continued with the Galileo Simulation. She began by leading the class in summarizing what they had learned during the first period. Whole class discussion then turned to the remaining activity sheet questions while a student operated the simulation.

About a half hour into the discussion, students began appearing bored and restless. The discussion lasted 53 minutes, spilling over into the next period by 10 minutes. In this class (Class A), Periods 2 and 3 of the lesson sequence occurred on the same day, separated by a ½ hour lunch break. The spillover occurred after the break, cutting into the time available for the Period 3 Projectile Animations lesson, as described below. Including the discussion from Period 1, total discussion time for the Galileo simulation activity sheet was 71 minutes.

Response to conceptual difficulties. An expression of confusion was noted in response to a question about the relative ranges and hang times for angles less than, equal to, or greater than 45 degrees. A student responded, "I don't know, I'm so confused." This issue was closely related to a topic of confusion on the previous day when the class had discovered that more than one launch angle could result in the same range. As before, the teacher did not change what she was doing to address the confusion, possibly because this phenomenon continued to be the topic of discussion without her intervention. There was also an episode where the teacher addressed a misconception that she knew many students have, even though it had not been in evidence in this class: the idea that the initial launching force of a projectile continues to act on it throughout the trajectory. Total time spent addressing conceptual difficulties and misconceptions was about five minutes.

Discussion about key concepts: Identifying concrete causal factors. Gravity was mentioned once and air resistance once. Interestingly, concrete causal factors were not invoked to address the above conceptual difficulty about ranges and hang times. Total time spent on these topics was less than ½ minute.

Period 3 (Analysis). The third period in the sequence will be analyzed in terms of Research Questions 2, 3, and 4. In this class, Period 3 was actually the last half of a double period. The discussion centered on use of the Projectile Animations (Video Clips

1-3). The teacher gave a two-minute introduction to the animations, then moved quickly to a whole class discussion with the activity sheet. After the class had discussed all the questions on this activity sheet, the teacher asked the students to give a summary of what they had learned. One student responded and the teacher concluded with her own short summary, recapitulating the student responses and confirming the conclusions.

Research Question #2: Discussion about key concepts: Identifying concrete causal factors. In this class, a fair amount of discussion about causal factors had occurred during the teacher's introduction in the first period of the lesson sequence. During this third period of the sequence, there was almost no additional discussion about concrete causes for projectile motion. Early in this discussion, a student remarked that the constant velocity shown by the horizontal velocity arrow in the Vectors Animation (Video Clip 1) indicated the absence of air resistance. The teacher repeated this comment in a confirmatory tone. This short exchange was coded as "discussion of a causal factor" even though the factor was not discussed further.

In another episode, the teacher asked students whether there is acceleration when a projectile reaches the top of its arc. A student replied that there is and the teacher asked why. The student replied that gravity is acting on the projectile. The teacher then asked whether gravity suddenly stops acting when the ball reaches the top of its trajectory and students agreed that it does not. This episode will be discussed in more detail below.

Total discussion about concrete causes during this class's use of the Projectile Animations and activity sheet was 20 seconds or 2% of discussion time.

Research Question #3: Response to conceptual difficulties and misconceptions. In the Period 3 discussion, there was one extended episode in which the teacher responded to several related misconceptions and another episode in which classroom discourse responded to the conceptual difficulty exhibited by a student.

The Vectors Animation (Video Clip 1) used animated arrows to represent the vector components of the velocity of a projectile. One student commented that the acceleration of the projectile was in the positive direction and another student responded that the acceleration was both upwards and downwards. Misconceptions about the direction of acceleration are common among high school physics students; actually, the acceleration due to gravity is downward at all points, even when the projectile is travelling upward. The teacher responded to the student misconceptions by asking questions to encourage discussion, and then suggested that it was as though someone were pushing on the projectile to make it slow down. She facilitated an extended discussion about the acceleration of an elevator on the way up, at the top, and on the way down. But when she asked about the direction of acceleration of a projectile when at the top of its trajectory, several students suggested that there would be no acceleration at this point. The teacher responded by asking probing questions, supporting students in interpreting visual aspects of the Vectors Animation, asking the students *why* acceleration was occurring in the first place, and restating student responses in a clearer fashion. She then encouraged the class toward consensus around the idea that, at the apex of the trajectory of the projectile, the vertical component of velocity is changing direction but the acceleration remains negative. One probing question that appeared to elicit correct student responses was, "Did the Earth stop pulling on the projectile when it was at the top of its arc?" This discussion lasted a little over two minutes. In the transcript of this episode, boldface indicates depictive gestures.

78 T: And **[G]** then once it reaches that **[G]** highest point, when it, uh, as it starts to come back down there, what direction is the acceleration then?
79 S: Down.
80 S: Positive.
81 S: It's still negative.
82 T: Sooo, what direction is it moving in?
83 S: Negative.
84 T: **[G]** Negative? Is it speeding up or slowing down?
85 S: Speeding up.
86 T: Speeding up. So **[G]** it's like it's being pushed down, so what direction would my acceleration be in?
87 S: South.
88 T: South. *(low laugh)* Still negative.
89 S: Down.
90 T: Still negative. All right, what about the very, very top?
91 S: Nothing.
92 S: Zero.
93 T: Let's think about this.

It is clear that some students thought that a projectile is not accelerating when at the top of its arc. A student then mentioned gravity as a cause of this motion and the teacher closed the discussion about the Vectors Animation by mentioning properties of vectors. She explained that at the top of the arc, the vertical velocity vector component, though it has zero magnitude, is changing direction. This discussion in response to misconceptions about the direction or existence of acceleration at the apex of a projectile's trajectory lasted a little over 2 minutes.

The expression of conceptual difficulty occurred a little later, during discussion about Lines Animation I (Video Clip 2), which used horizontal lines to show the progress of the projectile along the *y*-axis. These lines appeared at equal time intervals but at unequal spacing along the *y*-axis. The student exhibited confusion, asking, "What- What does the spacing mean?" The teacher responded with a *reflective toss* (Van Zee and Minstrell, 1997b), throwing the responsibility for reasoning back to the students by rephrasing the student question: "So what does the spacing, what is the spacing telling us?" Another student started to reply, then hesitated. The teacher remained silent. After several seconds S1 spoke up, using gestures and words to try to make the case that the spacing between the lines indicated that the velocity slowed on the way up, changed direction, and then increased on the way down. (This episode is discussed below in terms of visual support moves used by S1.) Some students continued to exhibit confusion and the teacher reminded the class of the way in which the spacing between the dots in a motion map represents velocity. The students appeared satisfied and the teacher moved on to another question. This discussion continued for slightly more than a minute.

Total discussion about conceptual difficulties and misconceptions lasted just under 3 minutes, or 18% of the discussion time.

Research Question #4: Support for key visual features: Lines, arrows, spaces, and acceleration indicators. There were many support episodes observed during this discussion, all but one of them by the teacher. The students appeared fairly animated and some of them were observed gesturing, but only one of the student episodes appeared designed to help fellow students interpret the visuals in the animation. This was an episode mentioned above, where S1 used gestures and words to describe the meaning of the vertical spacing between the horizontal lines in Animation II. In the transcript of this episode, square brackets indicate gestures, boldface indicates depictive gestures, and the underlined passage indicates the excerpt coded as visual support.

136 S: Wait. (*Pause*) What does the spacing mean?
137 T: So what does the spacing, what is the spacing telling us?
138 S: The first, um-
139 (*3 second silence, then a different student speaks*.)
140 S1: Telling you [*points toward display screen*] that it starts at a ***[points up along the beginning of an arc]*** higher speed, or [*points toward screen*] velocity, and slows down, and then as it [*points*] changes direction, it starts to ***[points to the side and begins to curve downward]*** speed up again.

141 T: Ok.

The student episode involved *selectively pointing out aspects* of the projectile's trajectory while *gesturing in the air to indicate those aspects as part of an apparent attempt to help other students interpret them*. His words, paired with his depictive and pointing gestures, helped indicate relationships between the line spacing, the vertical velocity component, and portions of the projectile's curved trajectory.

The teacher episodes involved *prompting questions*, *gestures*, and *hints to encourage students to interpret the meaning* of the velocity arrows and the spacing between lines in the animations. Fifty-three teacher support moves were identified during the discussion. Together with the student support move, this was an average of about 3 ½ support moves per minute or 207 per hour.

Other comments. The Period 3 whole class discussion about the Projectile Animations and animations activity sheet lasted 15 ½ minutes, all of which were videotaped and analyzed. This was only slightly less than the corresponding Period 3 discussions in the other two matched classes even though the Galileo Simulation discussion in this class had run long and spilled over into Period 3. Total time-on-task for the entire 3-lesson sequence was close to that of the other two matched classes.

ii. Whole Class Condition #2: Year One HP Teacher C

This was the second of Teacher C's three matched classes. Class B, like Class A above, was conducted in the whole class discussion format. The lesson in the first two periods involved use of the Galileo simulation and Simulation activity sheet while the lesson in the last period centered on use of the three Projectile Animations and the animations activity sheet. The first two periods will be summarized to give an idea of what kinds of student confusion had already been addressed and how much discussion about concrete causes for projectile motion had occurred before the students encountered the Projectile Animations. The third period, in which the animations and activity sheet were discussed, will then be analyzed in terms of the research questions.

Period 1 (Summary). The structure of the first period lesson was similar to that of Class A, discussed above. After the pre-test, the teacher led an introductory discussion during which she threw balls around the room and asked students what factors were affecting the motion. They answered with several concrete causes: applied force, gravitational force, air resistance. She then went over terminology to be used in the lesson. The teacher then handed out the prediction sheets and the students filled them out working independently. Next the teacher handed out the Galileo Simulation activity sheets and brought up the simulation, projecting it onto a screen in front of the class. The whole class discussion that followed lasted 17 minutes, until the bell rang.

Response to conceptual difficulties. There were no episodes of puzzlement, bewilderment, or frustration observed during this class period. This was Teacher C's second whole class discussion of the day on this topic and she altered some of her strategies—perhaps in response to student difficulties she had observed during the earlier class. One of these strategies involved mentioning concrete causal factors.

Discussion about key concepts: Identifying concrete causal factors. In this second class (Class B), when discussing what affect changing the mass of the simulated projectile would have on the trajectory, the teacher asked a question she had not asked Class A earlier. After explaining that the velocity and angle were to be held constant but that the mass was to be increased before the launch, she asked, "What am I actually gonna have to do, in order to get it to have that same initial velocity when it leaves my hand if I were throwing it?" A student responded, "You can throw harder." This was a question about concrete causes that the teacher had not asked of Class A. Interestingly, the students in Class B made incorrect predictions similar to those made in Class A, that increasing the mass would change the trajectory. However, when the trajectory remained the same, these students did not exhibit the confusion or frustration of the students in Class A, but offered explanations for what had occurred in terms of the difference in the applied force required to bring both masses to the same launch velocity. Discussion about concrete causes lasted about two minutes.

Period 2 (Summary). Due to the rotation of the class periods, the second part of the Galileo Simulation lesson was conducted in Class B before it was conducted in Classes A or C, so this Class B discussion was the teacher's first experience facilitating a discussion about the last part of the Galileo Simulation activity sheet. She began by reminding students what they had been discussing when the lesson had ended on the previous day. The class then discussed the remainder of the questions on the Galileo Simulation activity sheet, appearing fairly engaged throughout the class. The discussion lasted 40 minutes, finishing just as the bell rang. Including the discussion on the previous day, total discussion time for the Galileo Simulation activity sheet was 57 minutes.

Response to conceptual difficulties. As these same students had during Period 1, they again appeared interested, even intrigued, when results of the simulation did not agree with their predictions, but did not appear confused or frustrated. The only confusion occurred near the end of the discussion when there was a glitch in the simulation: at very large masses, the ending velocity read-out was greater than the initial velocity read-out. The students took some care to establish what part of the simulation was correct and what part incorrect, asking the teacher to clarify this.

Discussion about key concepts: Identifying concrete causal factors. The effect of air resistance was discussed, and gravity was mentioned once, but concrete causes were not emphasized as an explanation for phenomena during Period 2 as they had been during Period 1. In fact, discussion about concrete causes apart from direct questions on the activity sheet lasted less than ½ minute. To address the surprising fact that complementary launch angles produce the same range, students focused on establishing what quantities in the simulation read-out were the same for the two angles (*range*), and what were different (*greater height and greater hang time*). They appeared satisfied with identifying these observation patterns rather than with an extended exploration of why these patterns occurred. The one exception was when a student tried to explain the phenomenon by saying that when the velocity vector components were equal, their product would be larger than when they were unequal. This was not an explanation in terms of concrete causes but was an interesting formal argument.

Period 3 (Analysis). The third period in the sequence will be analyzed in terms of Research Questions 2, 3, and 4. During this period, the students used the three Projectile Animations (Video Clips 1-3) and the animations activity sheet. Because the class schedule rotated, on this day Class B met last, after the other two matched classes. Therefore, the teacher had already engaged in both whole class and small group discussions with the animations activity sheet and had some idea of the issues the students in this class were likely to face.

The class appeared to be in "couch potato mode," passive and relatively unengaged. Even though the teacher made many supporting moves, only a few students answered, and then in low monosyllables. No gesturing was observed among the students and it was not clear that all of them were paying attention. In general, the pacing was controlled by the teacher's rapid questions, while the student answers, though short, were prompt.

Research Question #2: Discussion about key concepts: Identifying concrete causal factors. During this last lesson of the sequence, discussion about concrete causal factors occurred twice, the first time initiated by the teacher and the second time by a student.

In the first episode, the teacher asked the students what forces act on a projectile. The ensuing series of teacher questions and student short answers established that 1) *gravity* acts in the vertical direction and so there is acceleration in that direction, 2) there are *no forces* acting in the horizontal direction and so no acceleration in that direction, and 3) the fact that the trajectory in the animation was indicated by a perfect parabola showed that there *was no air resistance* in the instance of a projectile portrayed there.

In the second episode, a student brought up gravity in answer to the teacher's question "How do we know (the acceleration is) in the negative direction?" The teacher was paraphrasing a activity sheet question intended to ask what *visual features in the animation* indicated the direction of acceleration, but the student answered the question in terms of his prior knowledge about gravitational forces. Rather than trying to clarify the meaning of the question, the teacher invited students to go more deeply into the issue raised by the student's response, the relationship between the downward force of gravity and the acceleration of the projectile. She used a scenario she had used in Class A, about an elevator. But where, in the earlier discussion, she had stated several times that it was *as though* the elevator were being pushed down, in Class B, in a subtle shift, she asked the students *in what direction one would have to push* on an upward moving elevator in order to slow it down.

109 T: Let's say that we have an elevator and it is **[G]** moving up, but it's slowing down. What direction would you have to push in, in order to get it to slow down?
110 S: Down.
111 T: It's **[G]** moving up, so we'd have to **[G]** push down on it? So we'd have a **[G]** downward acceleration? What if it started at the **[G]** top and **[G]** started to accelerate downward? What direction in that case would you be **[G]** pushing on it?
112 S: Down.
113 T: If I start at the **[G]** top, and I'm **[G]** moving down, I'd be **[G]** pushing downward to get it to **[G]** speed up? Does that make sense?

Thus, she invoked the action, the *push*, not as a metaphor, but as a concrete cause for the acceleration. This episode is an interesting example of support for student use of kinesthetic imagery (see Gooding, 1992; Clement, 2006); students were invited to imagine applying a force to produce acceleration in a given direction.

A moment later, when it became clear that not all students thought a projectile would undergo acceleration when at the top of its arc, she repeated a question about a concrete causal factor that she had also asked Class A, “Does the Earth suddenly stop pulling on the projectile when it reaches its maximum position?” Then she added something new—a thought experiment. She asked what would happen if she threw a ball up into the air, and suddenly at the top, the force due to gravity stopped working. A student responded that the ball would just stay there. Because that is not what we see happen, the class agreed that gravity was still acting. They then decided that what the acceleration of gravity was doing at that point was making the ball change directions.

Total discussion about concrete causal factors lasted 3 minutes, or 16% of the discussion time for this class period.

Research Question #3: Response to conceptual difficulties and misconceptions. The teacher addressed misconceptions about the direction of acceleration at the same point in the Projectile Animations lesson as she had in Class A. However, in the Class B episode about the elevator (Lines 109-113 above), she invoked causal factors to help with this, as she had not in the earlier discussion.

There was also one episode during this period in which a student appeared puzzled about a conceptual issue. The discussion had turned to Lines Animation II (Video Clip 3) which used vertical lines laid down at equal time intervals; these appeared at equal spacing along the *x*-axis to indicate the constant progression of the projectile along the *x*-axis. In the below excerpt, the teacher summarized the discussion up to that point, and a student responded with what appeared to be confusion. (Boldface indicates depictive gestures.)

150 T: There is no change in the velocity. In order to have an **[G]** acceleration, you'd have to have that spacing get larger or smaller. In this case, it **[G]** just stays the same, so our-

151 S1: *(sounding puzzled) What* stays the same?

The teacher responded to S1 by facilitating a deeper discussion about the meaning of the equal spacing between the lines in the animation. She asked what would have been the spacing between the vertical lines if the projectile had been speeding up in the horizontal direction (answer: spacing would have become farther apart toward the end of the range of the projectile) and what it would have been if the projectile had been slowing down (answer: spacing would have become closer together). She also invited students to think back to the spacing between dots in a motion map and to relate that to the spacing between the lines in the animation that was before them. The teacher appeared to be attempting to support students in making use of the visual indicators for acceleration (discussed more below) but may not have addressed an important underlying issue. S1 appeared to this observer to be having trouble conceiving of velocity and acceleration as divisible into separate and independent components, not an unusual difficulty for students at this physics level. However, rather than saying, for example, “This component of acceleration is zero,” the teacher made comments such as, “So, no acceleration here.”

The back-and-forth between teacher and students on this issue lasted 1 minute 23 seconds, or 7% of the discussion time during this class period.

Research Question #4: Support for key visual features: Lines, arrows, spaces, and acceleration indicators. There were many support episodes observed during this discussion, all of them by the teacher. The students had their backs to the camera and their gestures and words were not as visible on the videotape as were the teacher’s. However, during other discussions in that room in which the camera had been located in the same place in the classroom, student depictive gesturing had been observed.

The teacher’s visual support moves were similar to those she had used in the Class A discussion on this same activity sheet; she gestured frequently as she asked students to interpret features in the animations. She also gave a number of hints to encourage students to identify relationships in the animations as indicators for the presence or absence of acceleration. The following is an example:

21 T: Let's just play this again and just concentrate on the vertical arrow. So what happened to that arrow?

This single visual support episode was coded as *selectively pointing out a key feature in an apparent attempt to help students identify it as an indicator of the presence of acceleration* and as *asking a question to prompt identification of a relationship as an indicator of the presence of acceleration*. The teacher first selectively pointed out the vertical arrow, which indicated the component of velocity of the projectile in the vertical direction, then asked students a question to prompt them to notice that the length of the arrow changed as the projectile followed its trajectory. The relationships between the length of the arrow and the location of the projectile within its trajectory (the length changed) and between the arrow tip and its base (the tip moved downward relative to its base) were both important indicators of acceleration. The first relationship indicated the presence of acceleration in the vertical direction, and the second, its sign (negative).

There were 40 supporting moves identified during the 18 ½ minutes of videotape analyzed, or about 2 per minute.

Other comments. The Period 3 whole class discussion about the Projectile Animations and animations activity sheet lasted 18 ½ minutes. (A minute that was taken up with an unexpected interruption was not counted.) Total time-on-task for the entire 3-lesson sequence was close to that of the other two matched classes.

iii. Small Group Condition: Year One HP Teacher C

The third of Teacher C's three matched classes, Class C, was conducted in the small group discussion format. As with Classes A and B, the first two periods used the Galileo Simulation and simulation activity sheet while the last period used the three Projectile Animations and the animations activity sheet. The first two periods will be summarized to give an idea of what kinds of student confusion had already been addressed and how much discussion about concrete causes for projectile motion had occurred before the students encountered the Projectile Animations. The third period, in which the animations and activity sheet were discussed, will then be analyzed in terms of the research questions.

Period 1 (Summary). After the pre-test, the teacher led an introductory discussion during which she threw balls around the room and went over terminology to be used in the lesson. She then asked students what factors were affecting the motion. They mentioned air resistance, weight, force with which balls were thrown, and launch angle. The teacher mentioned force due to gravity and suggested that they consider the mass of the object rather than its weight. She then handed out the prediction sheets and the students filled them out working independently. Before handing out the Galileo Simulation activity sheets and sending students to their small groups, the teacher briefly brought up the simulation and showed students how to operate it.

This class had 13 minutes available for small group discussion on the first day and the small group on camera utilized 11 ½ minutes of it. Time available was several minutes less than for the matched Period 1 whole class discussions, though time available for the entire lesson sequence was similar. The camera joined a small group of three students. For narrative reasons, the group's Day 1 discussion about concrete causal factors will be summarized first.

Discussion about key concepts: Identifying concrete causal factors. While showing the students how to operate the simulation, the teacher repeated something that had arisen during one of the whole class discussions earlier that day, "If I'm throwing something that's heavier, then it's harder to throw, right? What am I gonna have to do in order to get an object that's twice the mass to have the same velocity when it leaves my hand?" When a student replied that she would have to use twice the force, the teacher invited students to imagine that the simulation was actually providing the extra force necessary to keep the velocity constant.

After the introduction, when the students had moved into their small groups, at least some of them appeared to remember this explanation. When the teacher stopped by the small group on camera, the students mentioned both the fact that the launching force would have to change for the greater mass and also that the greater mass would have more inertia. This discussion lasted only a few seconds, however, and the students in this small group did not discuss concrete causal factors except in connection with this issue.

Response to conceptual difficulties. There were no expressions of puzzlement or confusion observed in the small group on camera during Period 1 except occasionally in reaction to glitches in the simulation. When these students encountered the issue that had provoked so much curiosity and puzzlement in the Class A whole class discussion, and that had elicited whole class discussion about concrete causal factors in Class B, there was only momentary surprise in this small group, quickly resolved. The issue was that changing the mass of the projectile in the simulation did not produce a change in its trajectory. S3 responded that she had predicted this. S1 and S2 admitted that they had predicted something different, but then S2 immediately provided an explanation, "It did specify that it was the same everything else, just the mass changed, so the force changed, too." The student appeared to be referring to the introductory discussion about causal factors. The students appeared to be satisfied and the discussion immediately moved on to other topics.

Period 2 (Summary). The students went almost immediately into their small groups and resumed work on their activity sheets for the Galileo Simulation. The teacher stopped by the group twice and asked probing questions to motivate further thinking. The three students on camera finished the activity sheet questions in about 12 minutes and then explored the simulation for an additional 13 minutes, testing its limits and occasionally returning to issues on the activity sheet. For the last 4 minutes available to them, they engaged in off-task conversation.

Together with the discussion on the first day, total time available for Class C for the Galileo Simulation activity sheet was about 55 minutes, about the same length that had been available for Class B but less than that for Class A (which had run over into the following period with their discussion time). The small group on camera utilized about 37 minutes of the 55, although less than 24 minutes over the two periods was tightly focused on the activity sheet questions.

Response to conceptual difficulties. The students were surprised when more than one launch angle resulted in the same range. S2, especially, seemed frustrated and confused, "I give up. … What the hell? Do you guys have any inferences on our new data? Do you have any idea why it's doing that? It's weird." The students figured out that launch angles that gave the same range totaled 90 degrees but did not appear to know how to take their reasoning beyond this formal explanation. Discussion on this topic lasted about two and a half minutes.

Discussion about key concepts: Identifying concrete causal factors. There was a discussion about concrete causal factors when students returned to the issue of the effect of mass on trajectory. They decided that when air resistance was present, mass did become a factor even when the launch velocity was held constant. They observed that when air resistance was turned on in the simulation, the larger mass went farther. One student reasoned that, due to its greater inertia; air resistance would have less effect on the larger mass. This discussion lasted a little over a minute.

After the small group discussions, the teacher unexpectedly called the class back together and engaged the students in a 12-minute whole class wrap-up. With the simulation turned off, students reported answers they had arrived at for the activity sheet questions in their small groups. One student again mentioned the fact that the simulation appeared to change the launch force in order to produce a constant velocity for every mass. Another student mentioned that the greater inertia of a larger mass made it travel farther when air resistance was present.

Period 3 (Analysis). The third period in the sequence will be analyzed in terms of Research Questions 2, 3, and 4. During this period, the students looked at the three Projectile Animations (Video Clips 1-3) and used the animations activity sheet. The students in the small group joined by the camera went through the animations fairly quickly, but then the teacher came by and the students wrestled for a while in her presence about the interpretation of one aspect of the animations. This difficulty will be discussed more below.

Later, the teacher led an unexpectedly extended wrap-up discussion in whole class. As with the wrap-up on the day before, the teacher turned off the animations and encouraged students to report the answers their small groups had agreed upon for the activity sheet questions. The students then handed in their activity sheets. Although the comparison of interest in these qualitative comparisons is between whole class and small group discussions in conjunction with the Projectile Animations, some of the interesting strategies the teacher used after the animations but before the post-test will be mentioned briefly.

Research Question #2: Discussion about key concepts: Identifying concrete causal factors. There was no mention of concrete causes during this small group discussion as there had been on the previous two days when this same small group had been discussing the Galileo Simulation. However, during the wrap-up discussion after the animations were turned off, concrete causes were mentioned five times, twice by students and three times by the teacher. Four of these times were when the teacher reminded students of an elevator example she had used in this class during a prior lesson sequence. Students and teacher discussed what kind of force it would take to start and stop the motion of an elevator, using the acceleration of the elevator as an analog to acceleration due to gravity. (These episodes were not counted because they did not occur during the small group discussion that accompanied use of the animations; however, they are mentioned in the table of qualitative results with an asterisk. See Table 43 below.)

Research Question #3: Response to conceptual difficulties and misconceptions. The group on camera called the teacher over as soon as they experienced conceptual difficulty. The teacher stayed with them for about 3 minutes to support them in clearing up their difficulty in conceiving of a position-position graph that could also include information about time. At the beginning of the excerpt below, the students had just turned on Lines Animation II, which used a series of equally spaced vertical lines to represent the constant progress of the projectile along the *x*-axis (Video Clip 3). They were discussing an activity sheet question that asked which component of the velocity the lines gave information about. In the excerpt, square brackets indicate gestures and boldface indicates depictive gestures. The excerpt begins when the teacher arrived at the group.

90 T: How are you guys doing?
91 S2: Am I dumb or is this right? (*laughter*)

95 T: So what's happening to the velocity?
96 S2: Those lines have no bearing on the velocity itself. Those are just indicating the time. Not the position.
97 S1: Ahhhh-
98 S3: But, what does that tell you about the velocity?
99 T: How far does it go each second?
100 S2: There's no horizontal lines. It's-
101 T: What is that telling you about the velocity?
102 S1: Constant.
103 S2: No, it's not.
104 S1: Okay. (*Laughing*) How isn't it?
105 S2: Because this is the *[points to something on the display]* distance.
106 S1: Yeah? No, it's not. Yes, it is.
107 S2: Yeah, so there is a dot on each second.
108 T: What's being plotted on the *x*-axis?
109 S2: Time.
110 S1: Not necessarily because-

In Lines 111-116, S3 argued that the spaces between the dots indicated velocity, but S1 pointed out that they didn't know for sure that the *x*-axis was time. (These lines will be discussed in terms of visual support episodes in the next section below.)

117 S3: (*overlapping)* I say it's time; if I say it's time, it's time.
118 S2: (*overlapping*) The red lines [*pointing toward screen*] are only-
119 S1: (*overlapping*) Is it ***[with both hands, holds fingers at right angles to each other]*** distance and ***[shifts orientation of fingers]*** distance?
120 T: It is, in fact, distance and distance.
121 S2: How is it distance and distance?

131 S2: I thought- Wouldn't the *x*-position be time anyways? (*2 second pause*)
132 T: Can position and time be the same thing?
133 (*pause)*
134 S2 (*softly*): Noo. (*shakes head*)

By “Can position and time be the same thing?” the teacher appeared to be referring to intervals on an axis. If that is what she meant, then the answer would be yes, the representations for position and time can look the same *if* the change of position with respect to time is constant and the scale is chosen appropriately. However, this may not be the meaning that S2 took from her comment and it appeared to confuse him.

A moment after the transcript excerpt ends, S2 came back with the answer that, for the projectile, position vs. time would be the same graph as position vs. position. (This is correct except for a scale factor). By this time, S1 seemed to have accepted the teacher’s comment that it was a position vs. position graph, but he also seemed to misunderstand what this meant, insisting to one of his group mates, “There is no time involved.” In response, the teacher encouraged the students to think of the graphs as motion maps. After more discussion, the students appeared to come to an understanding of the position-position representation in the graph and the teacher left the small group.

Support moves the teacher used to address the difficulties experienced in this group included multiple visual support moves (discussed in the next section) that appeared designed not only to challenge the students’ interpretation of the visual representations in the graph, but to challenge assumptions that may have underlain their difficulty in interpreting the graph as position-position even after being told that it was. In addition to supporting students to arrive at their own understandings, the teacher also supplied parts of the answer at two points during the discussion, when she confirmed that the graph was, in fact, position-position, and at the end, when she encouraged the students to think of the representation as a motion map. Thus there appeared to be two phases in the teacher’s support: supporting students to open up their thinking and then supporting them to converge on the correct answer. (See Price and Clement, 2011, for an in depth exploration of this whole class discussion-leading strategy.)

Research Question #4: Support for key visual features: Lines, arrows, spaces, and acceleration indicators. There were a number of episodes when students supported each other in recognizing and interpreting the visual affordances of the animations. Most of these occurred while the teacher was present with the small group. The following excerpt is the missing portion of the extended excerpt under Research Question 3 above. (Lines 108-110 are repeated for continuity.) The teacher was responding to the students' difficulty in conceiving of a position-position graph that also included information about time. Square brackets indicate gestures, boldface indicates depictive gestures, and underlining indicates visual support moves. This transcript excerpt is further annotated to identify the visual support moves.

108 T: What's being plotted on the *x*-axis?

STRATEGY: Asks a question to prompt interpretation—or reinterpretation—of the meaning of the relationship between the vertical lines.

109 S2: Time.
110 S1: Not necessarily because-
111 S2: (*unintelligible*)
112 T: This (*the round dot travelling across the screen*) is just a ball going across the screen.

STRATEGY: Gives a hint to encourage reinterpretation of the spacing between the strobes.

113 S3: I say it's (*the x-axis is*) time. And I say this is [*moving eraser of her pencil a nearly vertical path across the screen*] distance.
114 S1: I don't say it's time, I just say that's how far it's ***[points away from himself]*** going.
115 S3: So if it traveled this far [*with fingers, brackets the distance between two strobes on the screen*], and then it traveled this far [*brackets the distance between the next two strobes to the right*], and then it traveled this far [*brackets next two strobes*], and then it traveled this far [*brackets next two*

strobes], it's traveling less every equal amount of time. So it's showing you the velocity-

STRATEGY: Gestures over the display to indicate the spacing between the strobes as part of an apparent attempt to help other students interpret their meaning, and also to identify the relationship between these distances as indicating that the velocity is changing.)

116 S1: (*overlapping*) -well, but we don't know that's time. It's like a ball ***[beginning in front of his chest, moves hand through an arc up and away from his chest]*** going- it could be ***[starting again from the same point, moves hand horizontally away from his chest]*** distance and ***[starting again from the same point, moves hand vertically upward]*** distance.

STRATEGY: Gestures in the air to indicate key vertical and horizontal relationships in the animation as part of an apparent attempt to help the other students re-interpret them as indicating distance vs. distance.)

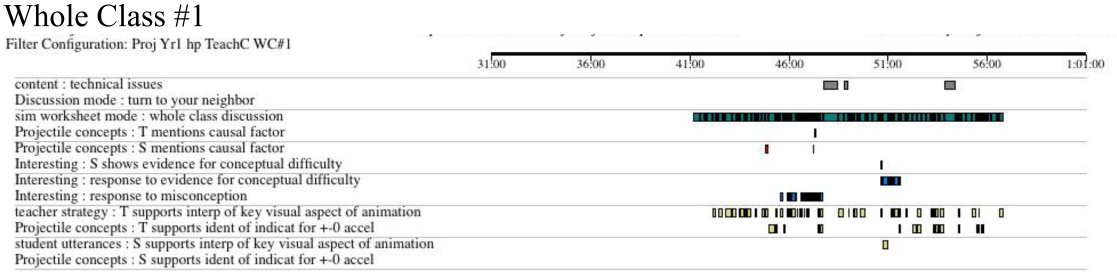
In this excerpt, the teacher made two visual support moves and the students made two. S1 and S3 used gestures as they tried to support each other to understand their differing interpretations of the relationships on the screen. Together with the excerpts in the previous section, the entire transcript section from Lines 90-134 included several visual support moves by both teacher and students. Lines 95, 104, 108 and 132 were prompting questions, Lines 99 and 107 involved selectively pointing out relationships, and Line 120, in addition to supplying information (not considered a support move), was also a hint about how to interpret features within the animations (considered a support move).

In all, there were 25 visual support episodes observed with 12 of them by the teacher and 13 by the students, at an average rate of 87 episodes per hour. Over half of the student episodes (and all of the teacher episodes), occurred during the 3 minutes when the teacher was present with the small group.

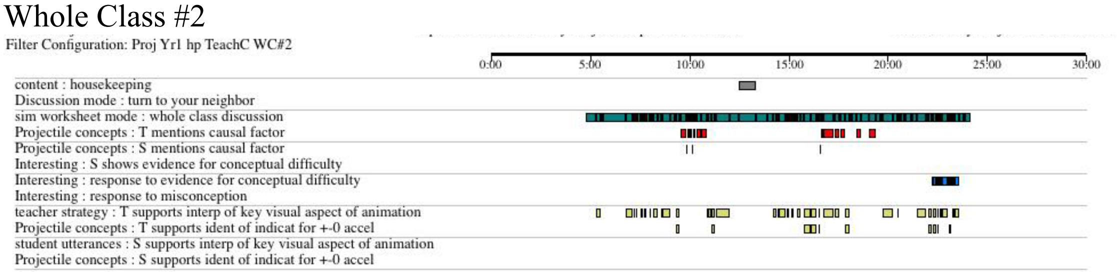
iv. Three-Way Comparison: Year One HP Teacher C

The videotape code maps (Figure 28) and Table 43 represent only Period 3 of the lesson sequence, during which the Projectile Animations were used. In the code maps, the transcript segments run chronologically from left to right. In the Small Group map, the unplanned follow-up discussion with animations off is also included in the interests of full disclosure; these codes are discussed in footnotes to Table 43. Color blocks below each transcript segment denote the codes assigned to that segment; the codes are listed on the left. The camera was used as a proxy for an individual student; the codes can be considered to reflect what an individual student in that class might have experienced.

Whole Class #1



Whole Class #2



Small Group

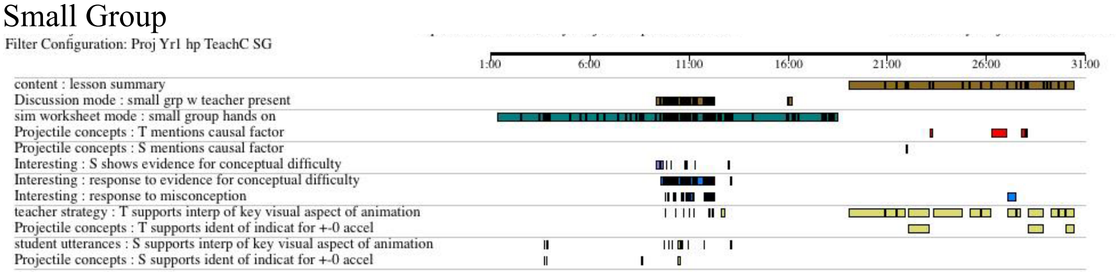


Figure 28: Videotape code maps: Year One HP Teacher C
(Each timeline spans 30 minutes of videotape, not all of which was taken up by classroom discussion.)

Table 43: Videotape coding results: Year One HP Teacher C

	A - Whole Class 1	B - Whole Class 2	C - Small Group
Time provided for activity sheets (Hand out until pick up)	17 ½ min	23 min	28 min*
Time provided for animations (including intro)	16 min	21 min	17 min 14 sec
Time utilized by students on camera for activity sheet questions (Starting at Q1)	15 min 37 sec	19 min 25 sec	17 min 14 sec
Technical & other difficulties	0	52 sec	0
Length of taped discussion analyzed below	15 min 37 sec	18 m 33 sec	17 min 14 sec
Research Q #2: Discussion about key concepts	Total length: **20 sec** Percentage of discussion: **2%**	Total length: **3 min** Percentage of discussion: **16%**	Total length: **0**** Percentage of discussion: **0%**
Research Q #3: Response to conceptual difficulties and misconceptions	Episodes of difficulty: 1 Response length: 1 min 3 sec Response to misc w no prior evidence of diff: Length: 1 min 50 s Total: **2 min 53 sec** Percentage of discussion: **18%**	Episodes of difficulty: 1 Response length: 1 min 23 sec Response to misc w no prior evidence of diff: 0 Length: 0 Total: **1 min 23 sec** Percentage of discussion: **7%**	Episodes of difficulty***: 8 Response length: 2 min 51 sec Response to misc w no prior evidence of diff: 0 Length: 0 Total: **2 min 51 sec** Percentage of discussion: **17%**
Research Q #4: Support for key visual features	Total support episodes: **54** Teacher: 53 Student: 1 Avg: **207 per hour**	Total support episodes: **40** Teacher: 40 Student: 0 Avg: **129 per hour**	Total support episodes***: **25** Teacher: 12 Student: 13 Avg: **87 per hour**

*The teacher unexpectedly led an 11-minute whole class wrap-up after the small group students had turned off the computers and returned to whole class. The students still had their activity sheets and had not yet taken the post-test. However, since presumption was in favor of the small group condition and this unexpected event favored that condition still further, it was of interest that this class did no better in pre-post gains than either of the two classes in the whole class condition.
** In whole class wrap-up, both students and teacher discussed concrete causal factors.

***In SG when teacher *not* present, there were only 6 visual support episodes and only 1 episode of conceptual difficulty, which had a 5 second response. The rest of these episodes were during the teacher's visit. In the unplanned whole class wrap-up with animations off, the entire eleven minutes was essentially an extended visual support episode; this wrap-up was not analyzed in detail but is noted as a possible factor in pre-post results.

Discussion. It can be seen from the table that the small group discussion had:

- **no discussion of concrete causal factors while in their small group;**
- many more episodes of conceptual difficulty than the two whole class discussions but comparable length of total response time to these difficulties;
- **considerably lower frequency of support episodes per minute than either whole class discussion** (about half the actual number of visual support episodes).

The most notable difference in the table above is in the amount of visual support. Only six visual support episodes were observed in the small group when the teacher was not present, although these classes, as a whole, were very rich in support episodes. These Projectile Animations class sessions were a challenge to compare because of the teacher's unexpectedly long wrap-up after the small group discussion. One thing that can be seen by looking at the code maps is that *most visual support episodes observed (including student episodes) occurred when a teacher was present*; i.e., during whole class discussion or when the teacher stopped by the small group. Even though the small group appeared to be a high functioning group compared to many of the others observed during the course of the project, over half the student episodes occurred during the 3 minutes the teacher was with their group.

In addition, *the only points at which discussion of concrete causal factors was observed on this day were during whole class discussion mode*. For the small group on camera, the only discussion about such factors came during the whole class wrap-up. The question arises whether this group had perhaps discussed concrete causes sufficiently during the preceding days while working with the Galileo Simulation that they did not need further discussion when working with the Projectile Animations in their small group. However, the transcripts for those days reveal that, although there was as much or more discussion about such causes across the three periods of this class as in Classes A and B, by far the majority (4/5) of this kind of discussion in Class C took place, not within the small group discussion, but during the whole class introduction and wrap-ups that the teacher led.

During the course of the project, this researcher and others came to suspect that a combination of whole class and small group work would yield the best results. The extended whole class wrap-ups in Class C, while not anticipated by the researcher, resulted in the unexpected chance to observe such a combination. The small groups had ample time on task with the worksheets and digital materials, comparable to that of the whole class discussions. In addition, the small groups had the benefit of a 12-minute whole class wrap-up of the Galileo Simulation activity sheet and an 11-minute wrap-up of the Projectile Animations activity sheet. These wrap-ups during Periods 2 and 3 were essentially extended visual support episodes in the absence of the digital visual tools, in which the teacher elicited responses from students and helped the class converge to a consensus on the worksheet questions. Therefore, although time on task with the digital tools (the simulation and animations) was similar in the two formats as taught by this teacher, total time on task with the *activity sheets* was considerably greater in the small group class. These students had the benefit of extended hands on work as well as the benefit of lively wrap-up discussions. Therefore, this comparison could most accurately be described as "Whole Class Discussion Only" vs. "Small Group Work + Whole Class Discussion." The inclusion of small group work and the additional time on task with the activity sheets would appear to have provided a substantial advantage to the Whole Class + Small Group Work condition. However, the short answer pre-post gains were very similar between the two conditions [$t(51) = 0.29$, $p = 0.77$, $d = 0.08$], and the effect sizes for those gains were large ($d = 1.43$ and 1.63 for WC and SG respectively). (In fact the gains were very similar among all three classes [$F(2, 50) = 0.06$, $p = 0.94$].) Even though interesting reasoning and active engagement was occurring in the small group on camera, the opportunity to engage in substantial small group hands-on exploration with the Galileo Simulation and Projectile Animations did not appear to confer a pre-post advantage on the small groups of Class C.

This result was also born out by the gains on the pre-post explanation questions. These questions concerned the effect on trajectory of changing the amount of gravity, the important causal factor in projectile motion. The gains were 9% and 24% for Whole Class 1 and Whole Class 2, and 12% for the Small Group Class. Thus, there is no evidence that the inclusion of hands-on activity conferred any advantage in reasoning about concrete causal factors to those in the Small Group condition. Though it might have been coincidental, the largest gains on these questions occurred in Whole Class 2, which also had the most discussion about concrete causes, though this discussion lasted only 3 minutes.

Support for the key visual features was rich in all three classes. Although support during use of the Projectile Animations appeared to be substantially greater in the whole class discussions than in the small group that was on camera, the Small Group class had the additional benefit of an 11-minute whole class wrap-up with the Animations activity sheet which, as mentioned, was essentially an extended visual support episode. Analysis of student written and drawn work in response to activity sheet explanation questions about the meaning of the animation features was used to provide an estimate of the extent to which these features were actually *used* by all of the students in the classes. Whole Class 2 and the Small Group class averaged almost the same on these questions, 68% and 66% correct, respectively. Whole Class 1, which had received the richest visual support during use of the animations, averaged highest on the activity sheet questions at 82% correct. Again, there was no evidence for an advantage for the Small Group class, even with the addition of the extended whole class wrap-up. This was one of the more surprising results in this study.

The fact that this teacher had two whole class discussion classes suggests an interesting comparison of a different sort. The opportunity to observe a teacher teaching the same lesson in the same format to matched classes on the same days was rare in this study. There were observed differences between Class A (Whole Class #1) and Class B (Whole Class #2) in terms of engagement, levels of confusion and pre-post gains. Some of these differences may be explainable by differences between students—the average pre-test scores were slightly higher in Class B than in Class A, and Class B had larger gains. A difference in execution was that in Class A, the teacher spent 10 minutes more on the Galileo Simulation activity sheet and 10 minutes less on the Projectile Animations activity sheet than she did in Class B. However, a potentially important additional factor is that the teacher seemed to respond to difficulties observed in the first class taught on each day and appeared to take steps to head off those difficulties in subsequent classes, principally by placing more emphasis on concrete causal factors. Even though the class order rotated[10], the lesson was taught first in Class A for two of the three lesson periods, and in both of those instances, the teacher spent less time on causal factors in that class. Over the entire lesson sequence, the teacher spent about 4x as long discussing concrete causal factors in Classes B and C as she did in Class A. Although it was not necessarily a direct consequence, it is interesting to note that the teacher spent at least 10 minutes addressing conceptual difficulties during the first two periods in Class A (the Galileo Simulation lesson), while no conceptual difficulty was observed in Class B during that time frame.

In closing, all three classes had very similar pre-post gains despite the differences discussed above. Even though they differed from each other along a number of dimensions, a striking feature of these classes, both during in-class observations and during video analysis, was the richness and frequency of the visual support episodes.

c. Year Two: Honors Physics (Teacher A)

This was the second year that Teacher A had taught the Projectile Motion sequence to his Honors Physics classes. (For Year 1, see the first Projectile case study comparison, Section C.3.a above.) As before, this teacher taught the lesson sequence as a two-period lesson; the periods were on subsequent days. The Projectile Animations were used during the second period. Each class's experience with the Galileo Simulation during the first period will be summarized to give an idea of what kinds of student confusion had already been addressed and how much discussion about concrete causes for projectile motion had occurred before the students encountered the animations. The Animations lesson will then be analyzed in terms of Research Questions 2, 3, and 4.

---

[10] The order of classes was: Lesson Period 1 - ABC; Lesson Period 2 - BCA; Lesson Period 3 - ACB.

The activity sheets were almost identical to those used the first year, the main difference being that the teacher removed a question from the Galileo Simulation activity sheet to shorten it.

i. Whole Class Condition: Year Two HP Teacher A

Period 1 (Summary). Before the pre-test, the teacher gave a brief introduction to the topic of projectiles and read a definition for projectile motion, "a body that is projected by external force and continues in motion by inertia." The class then completed the pre-test. Next, on the board, the teacher listed several terms the students would need for the lesson. While he discussed these terms, he tossed balls back and forth to the students. After the students had filled out their prediction sheets, the teacher turned to the Galileo Simulation and Simulation activity sheet and the class spent about 18 minutes in whole class discussion. After they had finished the activity sheet, the teacher continued another several minutes with the simulation, asking students about the effects of air resistance, the topic he had removed from the sheet in order to shorten it. The students seemed quiet and attentive but no marked reactions were observed except once when the students chuckled at a humorous comment.

Response to conceptual difficulties. No instances of student confusion were noted on the first day. In other classes, students had expressed surprise and confusion when increasing the mass of the simulated projectile had not resulted in a change in trajectory. In this class, when the teacher asked the students to predict what would happen when he doubled the mass of the projectile, two students correctly responded that the trajectory would remain the same because the velocity was constant. Another issue that had elicited surprise in classes was when shooting the projectile from complimentary launch angles resulted in the same range. In this class, the teacher did not ask for predictions beforehand but shot off the simulation at complimentary angles and then asked students to explain why they landed in the same spot. A student quietly responded that the angles were equidistant from 45 degrees and the rest of the class appeared satisfied with this. No specific responses to misconceptions were noted in this class.

Discussion about key concepts: Identifying concrete causal factors. Concrete causes were mentioned twice during the class but neither time was directly in connection with questions on the activity sheet. During his introduction to the class, the teacher had mentioned that an initial force launches a projectile. The second episode was when the teacher asked questions about the topic he had deleted from the activity sheet; he asked for predictions about the effect of air resistance and asked how to test these predictions. The discussion about this causal factor continued for a little over 3 minutes. If gravity, the central concrete causal factor in projectile motion, was mentioned at all, it did not occur during the videotaped portion of the lesson and was not noted in the observation notes of this researcher.

Period 2 (Analysis). The second period Projectile Animations lesson will be analyzed in terms of the research questions. As with many of the lessons observed, spontaneous events occurred in this class that were not anticipated in the lesson plan. After more than twenty minutes on task with the activity sheet and animations, insistent questioning from a student prompted the teacher to do something he hadn't planned; he performed a lecture demonstration that lasted almost 5 minutes. After a few more questions, the teacher began to wrap up the discussion. However, another series of student questions led to nine more minutes of discussion and mini-lecture. Finally, the teacher ended the discussion by saying, “My sense is, this is not helping for you. It's not clicking.”

The research questions are discussed here in the order 3, 4, and 2 to provide a chronological overview of the student activity.

Research Question #3: Response to conceptual difficulties and misconceptions. About 18 minutes into the discussion about the Projectile Animations, students began exhibiting confusion. They were trying to interpret the indicators for constant velocity used in Lines Animation II (a series of equally spaced vertical lines; see Video Clip 3) in light of the fact that they were positive there was acceleration occurring in the system.

87 T: Does anyone have a sense as to what equal distance in equal times indicates about velocity?
88 S: *(sounding puzzled)* I was gonna say that it was equal, but like, the velocity changes.

90 T: Can you say more about what it is that you see that's equal here?
91 S: No.

113 S: We're just confused because if it had a constant- if it didn't have acceleration, it would be linear.

A total of 14 ½ minutes, most of the remainder of the discussion, was coded as *response to conceptual difficulty* as the teacher guided the discussion to stay focused on the idea that both constant and accelerated components of velocity could be present within the same system. This represented an impressive 43% of the whole class discussion time.

Research Question #4: Support for key visual features: Lines, arrows, spaces, and acceleration indicators. There were 25 episodes with this code, all of them teacher moves. A particularly interesting episode occurred during the last part of the impromptu lecture-demonstration, almost a half-hour into the activity. Despite the fact that on the preceding day the students had appeared able to reason with the motion maps in the Galileo Simulation (see Figure 24), the teacher began to suspect from student comments such as those mentioned above that on this second day of the sequence, the students were misinterpreting the motion maps in the Projectile Animations (Video Clips 1-3) as *position-time* or *velocity-time* graphs. Though the students had worked with motion maps before, they were not accustomed to thinking of them as *position-position* graphs; the grid in the background of the animations appeared to confuse them. For a graph, they were accustomed to using the *x*-axis to represent *time* rather than *distance*. The teacher decided to try a lecture-demonstration although he did not have the equipment needed, and so had to ask students to imagine much of it.

After tossing a ball from the front to the back of the room several times, he asked students to imagine a spotlight shining from the back of the room toward the whiteboard on the front wall, casting a moving shadow of the projectile onto the whiteboard. He pointed out that, because the light was hitting the projectile head-on, this shadow would travel straight up and down. He drew marks on the whiteboard to represent the heights of the projectile's shadow at different points in time (Figure 29). He then asked the students to imagine another spotlight shining from the ceiling down to the floor and pointed out that the shadow of the projectile would travel across the floor.

The teacher then asked students to reason about this imaginary scenario (although it was not clear how many students had been able to follow his description). He turned back to the animation, projected onto a Smart Board positioned near the whiteboard (Figure 30). He used a marker to draw over the projected image of the animation, annotating the *y*-axis in the animation with marks. He pointed out the equivalence of

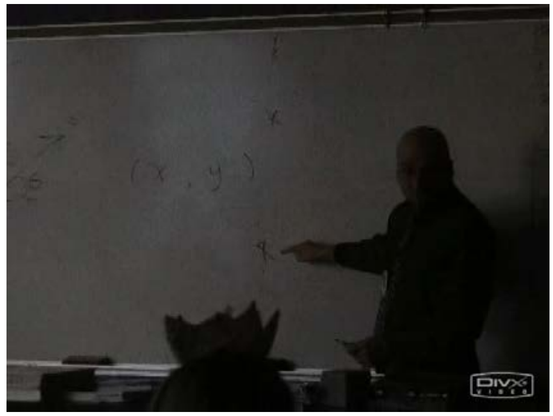


Figure 29: Teacher A indicating heights on the whiteboard, “The shadow goes up; the shadow goes down.”

these to the marks he had drawn earlier on the nearby whiteboard. He then drew marks along the *x*-axis in the animation and pointed out their equivalence to the shadows he had asked the students to imagine on the floor. Pivoting back and forth between his drawings on the whiteboard and the now annotated animation on the Smart Board, he described the equivalence in different ways, frequently adding to his annotations as he talked. This whole episode was accompanied by many depictive gestures as the teacher tried to demonstrate the components of motion, and sound effects as the teacher emphasized the equal time lapses between the appearances of the dots in the animated motion map.

The last minute and a half of this demonstration was coded in its entirety as a single long episode of *teacher supports interpretation of key visual relationships* and *teacher supports identification of indicators of the presence or absence of acceleration.* The visual relationships he was trying to help students interpret were the spacings between the vertical and horizontal lines in the animations. The teacher selectively pointing out these relationships with words and annotations, and he gestured both in the air and over the display to indicate them. (Although the teacher was working hard, it was not clear to what extent this lecture-demonstration was effective.)

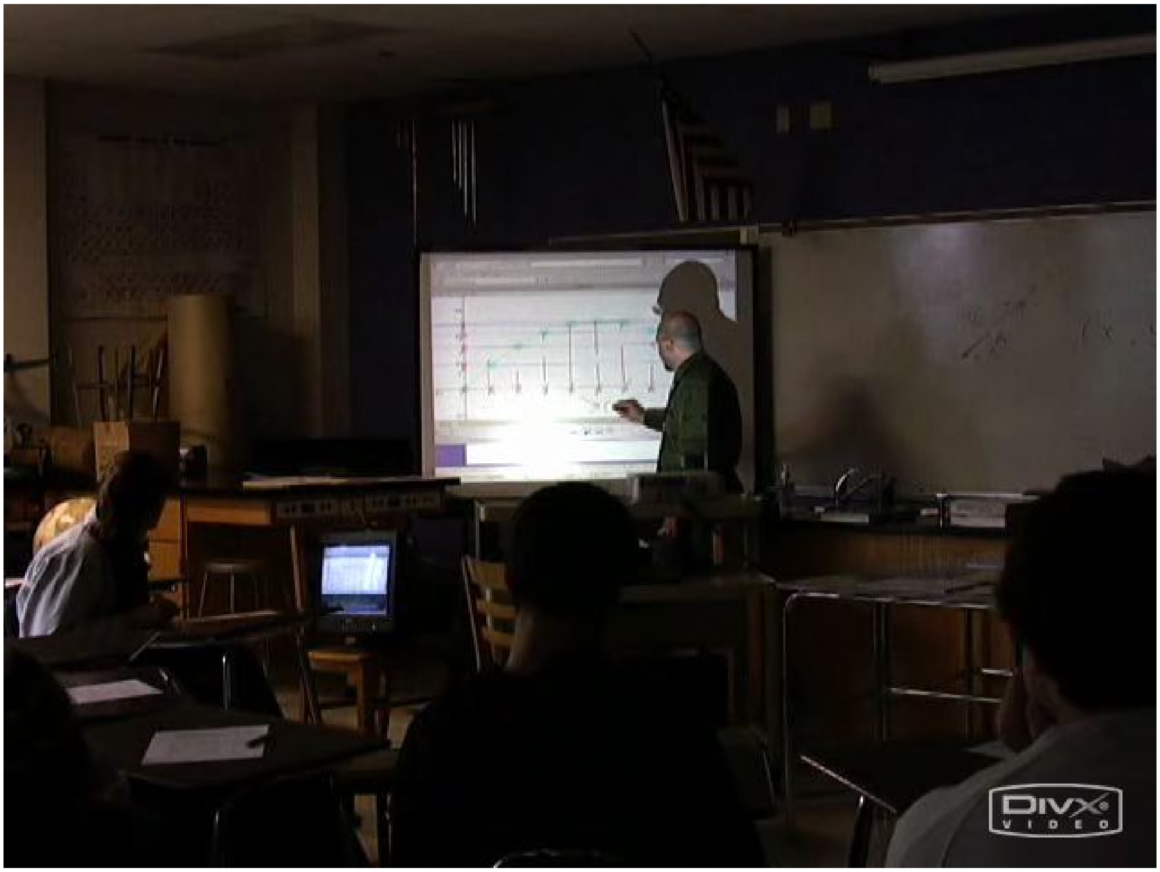


Figure 30: Teacher A Annotating the Animation on the Smart Board
Teacher pivoted back and forth between the Smart Board and the whiteboard to the right.

There were 25 visual support episodes for an average of 42 per hour of discussion.

Research Question #2: Discussion about key concepts: Identifying concrete causal factors. Gravity was mentioned only once in this class period, and then not in the context of a clear causal relationship. The following excerpt occurred near the end of the discussion, after the lecture demonstration.

139 T: And so this is really the essence of projectile motion. Motion with a constant horizontal velocity and a constant vertical acceleration.
140 S2: So wait, it's like acceleration and the velocity by the time, right? *(inaudible)*
141 T: I didn't quite follow the question.
142 S3: Wait, you just said there was constant acceleration, but I thought it was changing acceleration.

143 T: Acceleration is just gravity. *(Turning back to S3)* I actually didn't *(inaudible)*, didn't understand the question you asked.
144 S3: Well, like, if you try to find what the *(inaudible)*-
145 T: Define the time intervals between the dots?
146 S3: Yeah, the acceleration and the velocity?
147 T: Well, this movie, there's no units on these axes. The fact is, that most, I'd say 99.999 percent of the time, we deal with projectiles on Earth. And on Earth, we know the acceleration. It's 9.8, which we often round off to ten meters per second every second, in the direction down toward the local surface.

Gravity had not been discussed much, if at all, as a cause for projectile motion on the first day of the lesson sequence. It may be that it had been thoroughly discussed as a cause on an earlier day and that the teacher did not feel a need to discuss it further. However, on this day, rather than saying, for example, that the acceleration of a projectile is *caused* by gravity, or that gravity *produces* the difference between the characteristics of the vertical and horizontal velocity components, instead, the teacher stated that acceleration *is* gravity. For these students, many of whom appeared to have trouble distinguishing between velocity and acceleration, it would seem that this statement could have potential to increase confusion. Nonetheless, it was counted as discussion time in which *teacher mentions concrete causal factor* for acceleration. The episode lasted 37 seconds, or about 2% of the discussion time.

ii. Small Group Condition: Year Two HP Teacher A

Period 1 (Summary). The teacher began the small group class almost identically to the way he had begun the matched whole class discussion above. Before the pre-test, he gave a brief introduction to the topic of projectile motion, “I looked it up in the dictionary earlier today, and it says basically an object that is given a push and then allowed to continue on its own.” After the pre-test, the teacher continued his introduction to the topic by referring to the list of terms he had written on the board for the earlier class, tossing balls to the students to illustrate these terms. Students then filled out their prediction sheets. Next, the teacher showed the students how to navigate to the Galileo simulation and he briefly demonstrated the simulation controls. The students then spent between 16 and 19 minutes with the simulation in their small groups (depending on the group) and returned to their seats for a wrap-up. On this first day, the camera joined a group that had three students who appeared focused on the task at hand; they worked on the activity sheet for most of the time available to them.

Response to conceptual difficulties. Two episodes of student confusion were noted. In the first episode, the students were surprised that changing the mass of the projectile in the simulation did not change its trajectory. This led to about a minute and a half of lively discussion during which one of the students pointed out that it would be harder to throw a heavier mass at a given speed than a lighter mass, but that if one could, the result for the two masses would be the same. The students appeared satisfied with this reasoning and wrote it down. The second episode occurred when shooting the projectile at two different angles produced the same range. This appeared to stump the students for several minutes until one of them suddenly noticed that each pair of angles that produced the same range added up to 90 degrees. The students then decided that this was because the launch angles were equidistant from 45 degrees, the angle that produces the longest range. They did not reason about trade-off between hang time and horizontal velocity or discuss any concrete causal factors, but were satisfied with their explanation in terms of kinematics.

Discussion about key concepts: Identifying concrete causal factors. Concrete causal factors were mentioned in only one episode, when the students were reasoning about why changing the mass did not change the trajectory when speed and angle were held constant. Two causal factors were mentioned, gravity and launching force. One of the students argued that gravity would pull the heavier object down faster (incorrect), but another student said that that wouldn't matter because the question asked about objects already moving at the same speed. She then used the example of throwing an eraser vs. throwing her lab partner, saying that it would be more difficult to throw her lab partner at the same speed. The other two students (including the "thrown" lab partner) then agreed, laughing, that if the first student were able to achieve the same speed, the resulting trajectory for the lab partner and the eraser would be the same.

Lacking in this discussion was any reaction to the apparent misconception of one student that gravity would pull a heavier object faster than a lighter object; there was no discussion about the fact that the acceleration of gravity is constant. In fact, the act of considering the pull of gravity in connection with the question at hand appeared to be confusing for these students, while considering the difference in launching force appeared to satisfy them.

Period 2 (Analysis). The Projectile Animations and animations activity sheet were used during the second period. Again there were three students in the small group, two of whom had been in the group the first day. S1 was new, replacing a student who was absent. The group appeared to be relatively well functioning. The three students appeared to have a good background in the topic and may have been a little too advanced for the Period 2 activity sheet; at times they laughed over what they saw as the obviousness of the questions (though they were not always correct). They did encounter conceptual difficulties but appeared able to work through these together for the most part.

Although they were allowed 34 minutes for small group work, the group finished their animations activity sheet in less than 13 minutes and then turned to unrelated activities, unlike during Period 1 when they had utilized almost all the time available.

The research questions are discussed here in the order 3, 4, 2 to be consistent with the matched whole class discussion, above.

Research Question #3: Response to conceptual difficulties and misconceptions. There were many expressions of difficulty. Each time a student exhibited conceptual difficulty in this group, the discussion turned to that difficulty (if not already focused on it) and stayed with the issue until the students had agreed on an answer. This was not the case for some of the other small groups observed for this study, where issues were dropped before the group found a resolution. (See, for instance, sections B.3.a.ii. and C.3.e.ii. in this chapter.)

The excerpt below gives an idea of the joint reasoning about the animations that occurred in this small group. S1, who was new to the group, brought up something that had not been considered by the other two students the day before: the constant acceleration of gravity. The episode began when S3 indicated by words and gestures that she thought the acceleration was in the direction of the velocity, a documented misconception among a wide range of students (Angell, 2004). At this point, she did not appear to be aware of any conceptual difficulty, but to be seeking confirmation of the answer. S1 replied, "It's constant acceleration, due to gravity, right?" This was coded as *response to a misconception*. (It was also one of the few times in this small group discussion that a concrete causal factor was invoked as an explanation; see discussion about Research Question 2 below.) There followed about 4 minutes of discussion consisting entirely of a series of expressions of difficulty and responses. The first few seconds is given below. Lines 70 and 72 were considered *evidence for conceptual difficulty* and the remainder the *response* to this difficulty. Square brackets denote gestures and boldface denotes depictive gestures. The excerpt begins immediately after S1's comment that the acceleration is constant.

70 S3: ***[moves right hand as though throwing something vertically toward her partners]*** Projectile- is different than- wait, no.
71 S1: There is still a constant acceleration.
72 S3: *(looking at S2, slowly, questioningly)* Is it?
73 S1: Because- *(pause)*
74 S3: (*quickly*) Yeah, because, think about it, the acceleration, it starts, even though it starts a little ***[holds hand up, thumb and forefinger together, moves it in a short arc up and away from her face]*** negative, right, and it goes, it's still-
75 S1: So it's still *(inaudible)* down in *y*.
76 S3: *(overlapping)* If so, in what direction? *(pause)* In what direction, though? *(pause)*
77 S1: Well, it goes ***[points upward]*** both up and then ***[points downward]*** down.
78 S3: Yeah. What does it mean by direction, though?
79 S2: Like, is it going horizontal, is it-
80 S3: Oh, is it, acceleration in the-
81 S2: In the *y* or the *x*.
82 S3: Ohhh.
83 S1: *Y*, right?
84 S3: The *y*, the (*inaudible*) of the *y* is changing.
85 S1: Yeah.
86 S2: Yeah.

Even though S3 knew that the vertical component of velocity was changing and that the horizontal component was not changing, she still wondered whether the projectile wasn't accelerating along the direction of travel; her gestures helped clarify her question. S1 was sure that the acceleration was constant (due to gravity), saying that acceleration is "down in the *y*-direction," though a moment later he said "up and then down." S3 continued to express confusion about the direction of acceleration. Eventually, she and S2 agreed with S1 that the acceleration was in the vertical direction. A little later, they agreed that they knew this was true because the vertical velocity arrow was changing in length.

As S3 struggled to engage in complex spatial reasoning about line of travel and direction of acceleration, she made gestures depicting first the initial launching force and then the curved line of travel. S1, who was sure that the acceleration was in the vertical direction (though he thought the acceleration was up and then down), gave two very simple gestures, pointing in the vertical direction. While S3 was sorting through a number of variables, S1 was focusing on only one, the one he knew they needed for the answer, and his gestures likewise picked out this one variable.

Other than his first comment about gravity, S1 did not invoke concrete causes to address his partners' difficulties above. Rather, he gave declarative statements about the answer, while avoiding directly contradicting his partners and also allowing them time to reason. One question is whether it might have been his level of conviction, as much as his logic, that convinced S2 and S3.

Almost 5 minutes of this transcript was coded as *response to conceptual difficulty*. Although only about a third as long as in the matched whole class discussion, this was almost as great as a percentage of the discussion time, 38%.

Research Question #4: Support for key visual features: Lines, arrows, spaces, and acceleration indicators. There was one episode of *supporting identification of a feature as an indicator for the presence of acceleration.* S3 gestured to show S2 how the trajectory in the animation would look if there were no acceleration present in either the *x* or the *y* directions (Figure 31). Although both S1 and S2 appeared to understand her gestures plus words, it is doubtful whether they could have as easily understood her

Figure 31: Student gestures to provide visual support
S3 gestures to show shape of trajectory if no acceleration were present.

words alone, “Yeah but both of them, if both of them were- if it was going one way, the both ways, like constant speed, it would be a straight line.”

The students provided other kinds of visual support that did not address features of the animations. At one point, S1 drew on a piece of paper to show another student how a *time-velocity* graph would look. Although this did not fit the coding criteria because it did not directly address a visual feature of the animations, it seems likely to have offered indirect support for understanding the animations. The depictive gestures of S3 in lines 70 and 74 (in the excerpt above) also did not appear intended to help her partners interpret features of the animations, but to help her in imagining the physics in a real world situation. A possible explanation for the lack of support episodes for key visual features could be that these three students were fairly evenly matched in their understanding of the *visual elements* in the animations although S1 appeared to have greater *conceptual* understanding. Gesturing, for instance, appeared to be for purposes other than to communicate new information about the visuals to each other.

The single episode was equivalent to a rate of 5 episodes per hour.

Research Question #2: Discussion about key concepts: Identifying concrete causal factors. Gravity was mentioned three times as a concrete causal factor, twice by S1 and later by S2. As mentioned above, early in the discussion, S1 addressed an apparent misconception of S3 by replying that the acceleration was constant and due to gravity. S2 and S3 seemed to be comfortable with this statement, although they did not always appear to know how to incorporate this knowledge into their reasoning (as in Lines 75-82 above.) They did not mention inertia, the absence of a horizontal gravitational force, or any other possible reasons for the constant velocity in the horizontal direction, instead relying on observation patterns in the animations when reasoning about the horizontal component of motion.

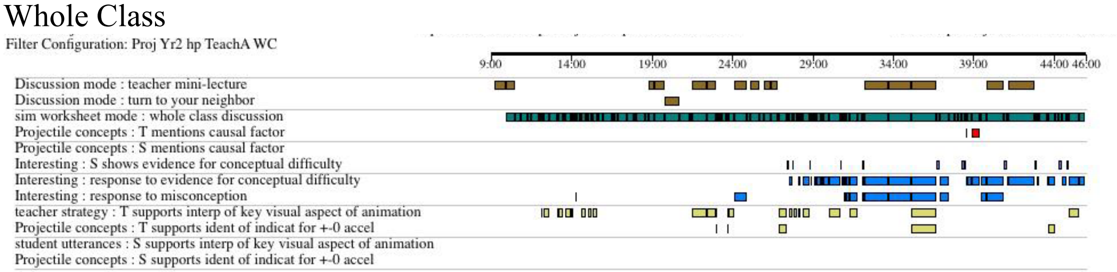
Total discussion about concrete causal factors was 39 seconds, or about 5% of the discussion time.

Other comments. By the time these three students had finished the activity sheet questions on the first of the three animations (Vectors Animation, Video Clip 1), they had figured out most of the concepts addressed by the entire Period 2 lesson: vertical acceleration, constant horizontal velocity, and the position-position nature of the graphs. However, it is not clear how deep of an understanding they had gained of the concrete causal factors that underlay their observation patterns. They finished the Animations Activity sheet in about 15 minutes although the class as a whole was given 34 minutes for this task and some of the other small groups utilized the entire time available.

iii. Comparison: Year Two HP Teacher A

The videotape code maps (Figure 32) and Table 44 represent only Period 2 of the lesson sequence, during which the three Projectile Animations were used. In the code maps, the transcript segments run chronologically from left to right, spanning the time when the students were working with the animations and animation worksheets. Color blocks below each transcript segment denote the codes assigned to that segment; the codes are listed on the left. In these videotapes, the camera was used as a proxy for an individual student; the codes can be considered to reflect what an individual student in that class might have experienced.

Whole Class



Small Group

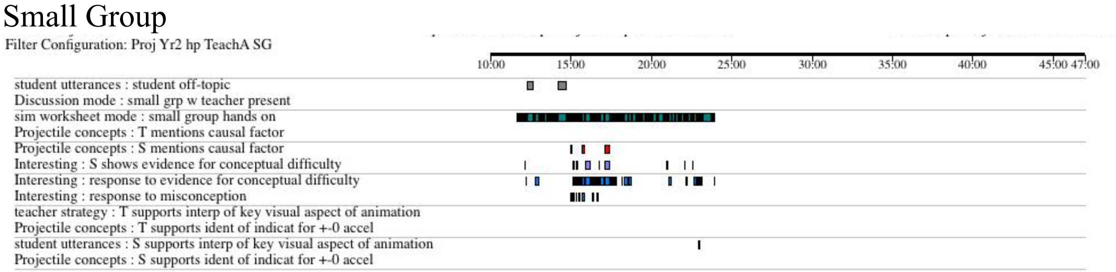


Figure 32: Videotape code maps: Year Two HP Teacher A
(Each timeline represents 37 minutes of videotape, not all of which was taken up by classroom discussion.)

Table 44: Videotape coding results: Year Two HP Teacher A

	Whole Class Format	Small Group Format
Time provided for activity sheets (Hand out until pick up)	39 min	39 min
Time provided for animations (including intro)	37 min	34 min

Time utilized by students on camera for activity sheet questions (Starting at Q1)	36 min	12 ½ min
Technical & other difficulties	29 sec	0
Length of taped discussion analyzed below	36 min 0 sec	12 min 25 sec*
Research Q #2: Discussion about key concepts	Total length: **37 s** Percentage of discussion: **2%**	Total length: **39 s** Percentage of discussion: **5%**
Research Q #3: Response to conceptual difficulties and misconceptions	Episodes of difficulty: many Response length: 14 min 39 sec Response to misc w no prior evidence of diff: many Length: 50 sec Total: **15 min 29 sec** Percentage of discussion: **43%**	Episodes of difficulty: many Response length: 4 min 40 sec Response to misc w no prior evidence of diff: 1 Length: 6 sec Total: **4 min 46 sec** Percentage of discussion: **38%**
Research Q #4: Support for key visual features	Total support episodes: **25** Teacher: 25 Student: 0 Avg: **42 per hour**	Total support episodes: **1** Teacher: 0 Student: 1 Avg: **5 per hour**

*The small group on camera took only 12 ½ minutes to finish their activity sheet, then moved on to unrelated work.

Discussion. It can be seen from the above that, compared to the whole class discussion, the small group on camera had:

- **1/8 the frequency of support for the recognition and interpretation of key visual features** of the animations (*1/25 the total number of support episodes*);
- **less than 1/3 the length of response to student difficulties**, although percentage of discussion time was similar.

Also, it is notable that there was:

- very little discussion about concrete causes in either discussion.

The small group attempted to provide each other with support for their conceptual difficulties and spent an unusually large percentage of their discussion time on this as compared with the other small groups analyzed in the study. However, there was a slightly higher percentage of response time to student conceptual difficulties in the whole class discussion and this occurred over a much longer period of time, resulting in several times the actual length of time spent on conceptual difficulties in the whole class situation. This comparison again points out a hazard of small group work already noted in case studies above—failing to take advantage of the entire time allotted. Although the small group students were fairly knowledgeable, they did not appear to know how to take their investigation with the animations any further than they did. It appeared common for small groups to turn to other tasks once they had finished the worksheet rather than delving more deeply into unresolved issues. In the whole class, on the other hand, the teacher filled out the time in an unplanned lecture demonstration that involved many visual support episodes, although it was not clear that the demonstration using annotation and imaginary shadows was effective.

It is interesting to compare these qualitative results to the results of this teacher's matched set during Year 1. He did not focus on *concrete causes* in his whole class discussions either year and the students in the two small groups on camera did not emphasize them either. A big change between Years 1 and 2 was in how much time the teacher spent on *addressing student difficulties* in the whole class setting; he almost tripled the percentage of discussion time he spent on this. In fact, an unplanned lecture demonstration the second year helped lengthen the discussion time beyond what he had planned. In raw minutes, the second year he spent over 6 times as long addressing student difficulties as he had the first year.

Even though the teacher spent considerable effort addressing conceptual difficulties and providing visual support, there was no significant difference between the pre-post gains on short answer questions in the two class formats and the effect size was negligible [$t(35) = 0.47$, $p = 0.64$. $d = 0.16$]. Both groups had significant gains at the $p < 0.001$ level with relatively large effect sizes ($d = 1.37$ and 1.23 for WC and SG respectively). On the explanation questions, the whole class students did appear to do better with 17% gains versus 6% for the small group, though gains were low and no statistical analysis was conducted on these results.

There was far more support for visual features observed in the whole class discussion, but there was joint reasoning about these features observed in the small group on camera. The students in all the small groups in the class averaged 75% correct responses on questions about these features on the activity sheets, compared with 74% correct in the whole class condition, suggesting that, over all, students in the two conditions were able to use the features to a similar degree in their reasoning.

At the time, the whole class discussion did not appear to have gone well and both the teacher and researcher believed that the small groups had done better. However, the videotape analysis, the activity sheet analysis, and the pre-post results indicate that the strengths and weaknesses of the two lesson formats appear to have balanced out. This result was similar to those of this same teacher's Honors classes the previous year. In fact, in both years, although the gains for the *short answer* questions on the pre-post test were similar for students in the two conditions, the whole class students appear to have had stronger gains on the *explanation* questions. Although no statistical analysis was done on the explanation questions, in each year, the raw gains for the whole class condition were double those of the small group students for this kind of pre-post question.

d. Year Two: College Preparatory Physics (Teacher A)

All of the projectile case study comparisons so far have been of Honors Physics classes, the mid-level group in the two schools represented. The next two case study comparisons will be of physics students at lower and higher levels respectively. This next comparison is of two matched classes of College Preparatory Physics (CP), a level that, though less advanced than Honors, was not a purely conceptual class; it utilized basic mathematics. Teacher A used the same materials as in his Honors classes, teaching the lesson as a two-period sequence with the Projectile Animations used in the second period. Each class's experience during the first period will be summarized to give an idea of what kinds of student confusion had already been addressed and how much discussion about concrete causes for projectile motion had occurred before the students encountered the Projectile Animations. The Animations lesson will then be analyzed in terms of Research Questions 2, 3, and 4. For these classes, the first and second periods were separated by a 2-day weekend.

i. Whole Class Condition: Year Two CP Teacher A

Period 1 (summary). The teacher's introduction to the lesson was very similar to the one he had given a week earlier in his Honors Physics classes (above). Before the pre-test, he gave a definition for projectile motion, "an object that has been given some motion and then it's allowed to continue on its own, and it only moves because of its own inertia." After the pre-test, he explained that in the Galileo Simulation, there would be three variables they could control: speed, angle, and mass, and three variables they would observe: maximum height, range, and time in the air. He gestured with two balls as he talked but did not throw them to the students. After the students had filled out their prediction sheets, they spent a little over 21 minutes discussing the Galileo Simulation and activity sheet in whole class discussion. The students seemed engaged and they participated actively and thoughtfully throughout the discussion.

Response to conceptual difficulties. There were two episodes of student surprise and confusion. The first was when increasing the mass in the simulation did not change the trajectory. Students gave exclamations such as, “This makes no sense!” However, one student quickly suggested a causal factor as an explanation, “Oh, 'cause if you're launching it at the same speed, you're putting more force on it, so like it's equal force, it's equal speed.” Other students replied, with some excitement and interest, that they had not thought of that.

The second episode of surprise occurred when two angles produced the same range. Again, a student quickly gave an explanation, although this time it was in terms of kinematics rather than in terms of concrete causal factors, “because they’re both 15 (degrees) away from 45.” Although several students then predicted that complimentary angles would produce the same range, not all students were satisfied. In fact, it took additional discussion and projectile firings at different angles before all the students were convinced that complimentary angles would work. Discussion concerned such factors as the differences in hang time and differences in total path length for different trajectories that had the same range; these concrete factors, though not causal, may have helped make the kinematic explanation more plausible.

Discussion about key concepts: Identifying concrete causal factors. The teacher did not mention any concrete causal factors in his introduction; in fact, the word “gravity” did not occur in this transcript at all. Air resistance was mentioned as a possible variable but was not mentioned in connection with what effects it might cause. Concrete causes were mentioned during only one episode, the episode described above in which a student introduced the topic of launching forces to explain why increasing the mass had not changed the trajectory. Causal factors remained the topic of the discussion for the next minute and a half as the teacher and several other students reiterated what this student had said, agreeing that additional force had been required to launch the larger mass.

Period 2 (analysis). The second period discussion will be analyzed in terms of the research questions. This discussion, about the Projectile Animations and animations activity sheet, was conducted on the Monday following the Friday lesson above. The teacher opened with a short review of the definition of projectile. He navigated to the three animations and spent time pointing out the features in the Vectors Animation both while it was on Pause and while he ran it. This amounted to a 3-minute mini-lecture during which he used gesture to give some visual support for interpreting features in the animations. He then turned to the activity sheet. After the class had worked through the activity sheet in whole class discussion mode, the teacher gave a brief wrap-up and then suggested that the students fill out the rest of their activity sheets. Students spent the last 2 minutes of the available discussion time writing quietly, although they were free to confer with each other if they wished. The length of whole class discussion accompanying work with the activity sheets was a little over 18 minutes. For narrative reasons, the research questions will be discussed in the order 2, 4, 3.

Research Question #2: Discussion about key concepts: Identifying concrete causal factors. Although concrete causal factors had been discussed during only one episode the first period, during the second period lesson about the animations, they were mentioned six times, the first three times brought up by students and the last three by the teacher.

Early in the discussion a student mentioned "wind resistance." A little later, when discussing the Vectors Animation (Video Clip 1), a student asked, "If there was wind resistance, what would the horizontal (velocity) arrow look like?" A second student suggested that it would cause the horizontal arrow to shrink unless "the projectile was so light that it would start to move the other way." While it is unclear whether students were confusing the *air resistance* from the previous lesson's Galileo Simulation with "wind resistance" due to moving air, it seems clear that at this point they were thinking in terms of moving air. The teacher did not try to disambiguate terms, but asked students what would happen to the path of the projectile if there were a strong wind present, "would it be a parabola like this?" The students decided that, even though it would still look similar to a parabola, it would no longer be symmetrical.

Later, a student mentioned gravity as the cause of the decrease in velocity as a projectile travels upward. The teacher repeated this comment so that the whole class could hear and then referred back to it again later in the class. Just before he wrapped up the lesson, he asked, "What causes that acceleration?" and two students answered, "gravity."

The other "causal factor" (or, more accurately, lack of a cause) that was mentioned was inertia. When the students were discussing the fact that a projectile undergoes no acceleration in the horizontal direction, the teacher commented that there was no force on it in the horizontal direction and that it travels in that direction solely due to inertia. Whereupon, a student astutely commented, “Gravity isn’t horizontal.” Equating the lack of gravity with the lack of a cause may have been easier for these students than reasoning about inertia, a difficult concept for many high school physics students.

There was a total of one and one half minutes of discussion about key concepts, or 8% of the discussion time.

Research Question #4: Support for key visual features: Lines, arrows, spaces, and acceleration indicators. The teacher engaged in three visual support moves during his introduction to the animations, but only the moves that occurred during use of the animations to address the activity sheet questions are considered for comparison purposes. During discussion about the Vectors Animation and Lines Animation I, there were 15 episodes in which the teacher supported the identification and interpretation of key visual features of the animations. Many of these support episodes included depictive gestures. However, once the teacher brought up Lines Animation II (Video Clip 3), only one additional visual support move was observed, a prompting question in Line 128 (see below) *to identify relationships* within the animation *as indicators of a lack of acceleration* in the horizontal direction. It is interesting that he provided no additional support for interpreting the visuals because in this part of the discussion he was addressing a persistent misconception about acceleration, something the visuals were designed to help. However, as can be seen in the extended episode below, the teacher used many depictive gestures, some of them in connection with the animation visuals. A plausible explanation is that by this time in the class, rather than trying to *support* *understanding of the visuals*, he was relying on the visuals to help *support understanding of the physics*.

The students were not observed providing any visual support for the animations during this class period though they used depictive gestures for other purposes on at least ten occasions, generally to help describe their thinking. The teacher engaged in an average rate of 49 visual support moves per hour.

Research Question #3: Response to conceptual difficulties and misconceptions. Unlike during the previous period with the Galileo Simulation, during this period students gave no evidence for being aware of or concerned about conceptual difficulties. However, there were two episodes when misconceptions about acceleration were in evidence to the observer, if not to the students. Even though the discussion responded only to the second episode, and therefore only this episode was coded "response to a misconception," it is of interest to look at both episodes because they were related.

In the first episode, a student stated that the acceleration for a projectile is constant (this is correct) because the *change in velocity* is symmetrical on the way up and on the way down and is zero at the top. Actually, the *velocity* is symmetrical on the way up and on the way down, with one of its components zero at the top, while the *change in velocity* is constant and equal to *g*. Although the student utterance could have been a simple misstatement, a common student misconception is that the acceleration of a projectile changes sign at the top and is zero at the apex of the trajectory. The teacher did not respond to the mistaken utterance, but only to the correct part of the student's comment. (The teachers reported that they sometimes chose not to respond to misconceptions in order to keep a lesson on course.)

However, late in the lesson, when S1 indicated (Lines 129, 131, 133) that a projectile will accelerate in the direction in which it is going, “to the right,” the teacher did spend some time on this misconception, engaging the class in discussion about it for around a minute. In the extended excerpt below, boldface indicates depictive gestures. Line 128, underlined, is the visual support move that was mentioned above.

127 S1: Say you're in space, right? That wouldn't be the same, obviously. Would it like, I don't know, would it still have a ***[uses pencil to trace back and forth along an arc in the air]*** bit of a parabola? Or would it just, I don't know, go straight?

128 T: Well, let's think about it for a second. Does this simulation show acceleration, and if so, in what direction?

This question prompted students to identify relationships in Animation II that indicated an absence of acceleration in the horizontal component of velocity.

129 S1: Uhh, ***[thrusts hand horizontally forward]*** that direction. The direction it's going in.

130 T: Acceleration ***[palm down, hand flat, thrusts hand horizontally forward]*** this way?

131 S1: Yeah. Or whichever way it's going. It's- going to the right. Yeah.

132 T: OK, so it's ***[moves hand horizontally back and forth across the display screen, repeats several times]*** accelerating to the right?

133 S1: Yeah.

134 T: It is?

The teacher took care to make sure he understood the exact nature of this student’s alternate conception before continuing. The existence of this misconception at this point in the lesson was surprising because students had been giving correct answers about acceleration all during the class. This episode raises the question of how much students with strong misconceptions had gotten out of the discussion—even though the explanations by their classmates had been lucid and even though there had been no debate about those explanations at the time.

Two other students chimed in with the correct answer:

135 S2: No. It's not.
136 S3: Accelerating down.
137 S1: Oh yeah.

The teacher continued with more explanation:

138 T: So, it's ***[starting with hands shoulder height and in front of himself, points downward with both hands and moves them quickly downward]*** accelerating down. That's what gives it its ***[moves hand across display screen along the line of trajectory]*** parabolic shape. What causes that acceleration?
139 S1: Gravity.
140 S4: Gravity.
141 T: Gravity, so, if we were out in space, meaning, OK, we can't measure any gravity out here, then would it go ***[points down and moves hand down suddenly]*** down?
142 S1: No.

This episode illustrates how persistent such misconceptions can be and suggests that Teacher A and Teacher C had good reason to elect to spend an entire lesson on the three apparently simple animations.

Total time spent addressing student conceptual difficulties and misconceptions was about a minute and a quarter, or 7% of the discussion time.

ii. Small Group Condition: Year Two CP Teacher A

Period 1 (Summary). As in the matched class, the teacher's introduction was very similar to the one he had given in his Honors Physics classes (above). He fleshed out his definition of projectile slightly, saying, "a projectile describes an object whose motion is started by something pushing it, but then it's just allowed to travel on its own through the air, and the only thing that's influencing it's just whatever is happening in nature. So after it leaves my hand, that stress ball is a projectile, in between when it leaves my hand until it reaches 'Steve's' hand, OK?" After the pre-test, the teacher tossed balls to the students to demonstrate the variables that they would be manipulating in the simulation. After the students filled out their prediction sheets, the teacher brought up the Galileo Simulation for about two minutes and showed all of its features, firing the simulated projectile once. Students then moved into their small groups, where they had about 17 minutes available to work with the simulation. The small group on camera finished the activity sheet in 12 minutes and then began to experiment with the simulation. The teacher came by and suggested that they explore air resistance, a topic not on their activity sheet, and they spent another couple of minutes doing this with the teacher present. The four students on camera did not always seem attentive and the one who controlled the mouse did not always appear to be taking the activity seriously.

Response to conceptual difficulties. There was one episode of conceptual difficulty in this small group when changing the mass did not change the trajectory. The students responded with comments such as, "I don't believe that," "It's false," "That doesn't make sense." One student replied by saying, "It does, because everything falls at the same distance." This student may have been trying to say that everything falls at the same *rate*. However, during follow-up interviews, when reporting on a class demonstration that had been conducted shortly before the Projectile Lesson Sequence, students made such comments as, "Everything takes the same amount of time to fall" or "Angle doesn't matter because the teacher showed us—I know it doesn't make sense." Although it was not clear to the researcher what this student meant, the other students did not follow up on his statement, but laughed and made comments such as, "Yes, the mass doesn't matter." They wrote down their answers and turned to the next question without further discussion on this issue.

Discussion about key concepts: Identifying concrete causal factors. When the teacher was introducing the topic of projectiles, defining terms and demonstrating them with a ball, he demonstrated what would happen if he varied the force at which he threw the ball and how that would change its trajectory. He did not mention gravity; in fact, gravity was not mentioned at all on the videotape. After the students had finished the activity sheet and were exploring the simulation, the teacher stopped by their group and asked them about air resistance. He gave them several suggestions for exploring this and supported them as they discussed air resistance as a causal factor for the changes in the shape of the trajectory.

Period 2 (analysis). The second period discussion will be analyzed in terms of the research questions. As in the matched class, the Projectile Animations were used on a Monday following the Friday lesson with the Galileo Simulation. The teacher opened with a short review of the definition of projectile. He navigated to the three animations and spent time pointing out the controls; this amounted to a 2-minute mini-lecture. However, he did not point out any of the visual features within the animations, comment on the strobes, or try in any other way to support students' interpretations of the animations before they broke up into small groups. Once the students arrived in their small groups, they had about 19 minutes available to work on the animations activity sheet and the small group on camera utilized all of this.

The small group was composed of the two female students from the on-camera group from the previous Friday. The two male students were not present on this day.

The teacher stopped by about halfway through the small group discussion time and then remained with this group throughout the remainder of their discussion, about 9 minutes. Although an individual small group is not likely to get a teacher's complete attention for half of their discussion very often, there are some possible reasons why he may have chosen to do this on this occasion; this group appeared to be having a fair amount of conceptual trouble. This group was not one of the better functioning small groups but was not the least functioning either; they worked steadily on the activity sheet with almost no off-topic discussion and they responded to each other's questions. S1 spent 19 minutes working with the activity sheet, all of the time available. S2 was away from the table for a little more than two minutes at the beginning of the small group time and spent about 17 minutes on the activity sheet.

The research questions will be discussed in the order 2, 4, 3 for narrative reasons.

Research Question #2: Discussion about key concepts: Identifying concrete causal factors. About half way through the discussion time, the students called the teacher over to their group because they were having trouble. Up to this point they had not mentioned any concrete causes for the behavior of the projectile. Once the teacher arrived, he asked them several questions about what was causing the acceleration and they mentioned gravity numerous times. Immediately before the episode below, S1 had said, correctly, that the horizontal velocity of a projectile does not change but that the vertical velocity does. In the transcript excerpt, square brackets indicate gestures and boldface indicates depictive gestures.

183 T: What do you think is making it *(the vertical velocity)* change?
184 S2: The hor- uhh-
185 T: For a real projectile, what makes that, whatever it is that [*points to "projectile" on display screen*] this thing represents, what makes-

186 S2: Gravity?
187 T: -it change?
188 S1: Or, it's running out of- ***[With hand cupped, raises it as though lifting something]***
189 S2: Gravity?
190 T: It's running out of something, whatever it was that it had at the beginning, OK?
191 S1: Momentum?
192 T: And, you're using the word gravity to describe-
193 S1: 'Cause gravity is pulling it back down.
194 T: -gravity is really just- OK, gravity pulls on it. So when something is going up ***[hand moves up to head height]***, and gravity pulls on it, that makes it-
195 S1: That's why it's slowing down.
196 T: And then when something is going down ***[slowly turns hand until it is moving horizontally, then downward]***, gravity pulls on it-
197 S1: It's pulling even faster, so it gets faster and faster and faster.

A little later, the teacher again invoked gravity as a causal factor as he probed these students' understanding of the nature of the acceleration as a projectile travels upward.

Total time spent discussing causal factors was slightly less than a minute, or 5% of the discussion time.

Research Question #4: Support for key visual features: Lines, arrows, spaces, and acceleration indicators. Teacher A elected not to discuss the visual features during his introduction of the animations, explaining to the students, "I don't want to give too much away because I do want you to investigate, think about, and discuss with each other, what do we think is happening?" He focused his introduction, instead, on giving a thorough demonstration of the QuickTime© controls.

During the first half of the small group discussion on camera, while the teacher was not present, there were six episodes in which one of the two students attempted to support the identification of an indicator for the presence or absence of acceleration, although these were not always correct. An example of student support that appeared to be successful occurred after S2 said she did not think the Vectors Animation (Video Clip

1) showed acceleration. In response, S1 asked, “You think that when it was going this way [*moving the curser to indicate a portion of the trajectory*], it was going the same speed as when it was going here [*indicating another portion of the trajectory*]?” She then paused and waited for S2 to answer. In this visual support episode, although S1 did not point out the intended indicator for acceleration (the movement of the vertical arrow), she *pointed out a visual relationship that indicated the presence of acceleration*, in this case, the difference in length of two segments of the trajectory. She also asked S2 a *prompting question*, prompting her to compare the two visual elements.

The teacher began his visit with the small group with a series of visual support moves that appeared intended to help students interpret key visual relationships in the animations. An interesting exchange involving several support moves occurred when the teacher prompted S1 to draw on a piece of paper in order to help her interpret a visual relationship in the animation. To help her see that *horizontal* lines can mark off a *vertical* distance, the teacher asked her how she would find the *y*-value of a given point on a Cartesian graph (a *prompting question* and a *hint to encourage interpretation of the meaning of the horizontal lines* in the animation). She showed him by drawing on a piece of paper and he responded, “Aha! You drew a line that way,” pointing to the horizontal line she had drawn from the point in question to the *y*-axis. Thus, he *pointed out a relationship* in her drawing that was the same as *a key relationship involving the horizontal lines* in Lines Animation I. The teacher next asked S1 about the “progress” the projectile was making in the *x*-direction and in the *y*-direction, a *prompting question supported by gestures*. She was able to say that it was moving at a constant rate in the horizontal direction. He suggested that the students bring up Lines Animation I and asked S1 another *prompting question*, “Which axis is this telling us about?” She replied, “Vertical.” The discussion continued with additional support moves, <u>underlined</u>.

152 T: And <u>what do we notice about the movement on that axis?</u>
153 S1: Not equal.
154 T: So, if I ask you, "Is the object accelerating along this axis?"
155 S2: Yes.
156 T: Just to be clear, <u>what is it about the motion that allows you to say that with confidence, "It's accelerating?" What do you notice?</u>

In Line 152, the teacher *asked a question to prompt interpretation of the meaning of a key relationship* (unequal movement in equal times in the *y* direction) and in Line 156, *asked a question to prompt identification of a key relationship as an indicator of the presence of acceleration.*

At this point, about four minutes into his time with the group, the teacher began to focus on helping S1 identify and interpret visual indicators for the presence of acceleration. After several prompting questions about the variable spacing between the lines in Lines Animation I, with the students referring to the animation and gesturing over it, S1 seemed clear on the idea that the projectile would go slower and slower as it rose and then faster and faster as it came back down. The teacher then asked the students to bring the Vectors Animation back up and asked what they noticed about the animated arrows (which represented the horizontal and vertical components of the velocity). S1 and S2 agreed that the horizontal velocity did not change but that the vertical velocity did.

Finally the teacher returned to the most challenging issue, about the direction of the acceleration. The last three minutes of the discussion between teacher and students focused on determining the direction of the acceleration while the projectile was rising.

Excerpts from this part of the discussion will be discussed below in terms of the teacher's responses to the students' conceptual difficulties.

The students engaged in 6 visual support episodes, a rate of 19 per hour. The teacher engaged in 17 episodes during his time with the group. Together, there was a rate of 72 support episodes per hour (teacher plus student) for this small group.

Research Question #3: Response to conceptual difficulties and misconceptions. Before the teacher joined the group, there had been three instances where one of the students responded to a perceived misconception or conceptual difficulty of the other. The first episode resulted in about a minute of discussion between the two students about acceleration. In the second episode, also about acceleration, S1 exhibited doubt and confusion and S2 responded merely by reading what she had written for an answer; she made no attempt at an explanation. In the third episode, S1 said that it didn't make sense to her that Lines Animation II would represent horizontal velocity because the lines in the animation weren't horizontal. In response, S2 stated simply, "The horizontal velocity is constant. The horizontal velocity is constant."

Once the teacher joined the group, the entirety of his activity appeared to be in direct response to the conceptual difficulties of the two students. This lasted for more than nine minutes. The teacher's efforts to address these difficulties were particularly interesting because he appeared to find it something of a challenge to think of strategies. At one point after a pause, he explained, "I'm pondering what kinds of questions I can ask you that will cause you to have some meaningful thought." The students found this amusing.

In the last three minutes of discussion, the teacher attempted to address the particularly difficult issue of the direction in which a projectile accelerates while it is rising. Many students find it counterintuitive that the acceleration is downward (due to gravity) while the object is rising. The teacher knew this and spent some time trying to get the two students to reason about this concept. While S2 played with the animation, pulling the slider back and forth to manipulate it manually, S1 exhibited frustration that the teacher wouldn't just tell her in what direction the acceleration was. In the following excerpts, boldface indicates depictive gestures.

205 S1: So, so, just to recap, what direction is the acceleration?
206 (*S2 laughs.)*
207 S2: (*laughing, collapsing on the lab table*) We never figured that out yet! There's no (*inaudible*) to figure it out!
208 T: Well, is it side to side?
209 S2: No.
210 T: What's the alternative, then?
211 S1 and S2 (*together*): Up and down!
212 T: All right, there you go.
213 S1 (*sounding incredulous*): So that's what we say? "Yes, up and down?"
214 T: Well, is it accelerating ***[moves hand upward]*** up?
215 S1: It's not accelerating, it's decelerating up.

S1 preferred to call what was happening on the way up "deceleration" although she admitted that she knew she was supposed to refer to this as "acceleration." "Negative acceleration" did not appear to be a part of her vocabulary; if not, that may have been part of her problem. The teacher attempted to address the issue.

230 T: You said (*pointing to S1's activity sheet*) "up and down," but-
231 S1: We should just say "up"?
232 T: -somebody else would then ask you, well, does it accelerate ***[with a loose fist, thumb pointing upward, moves hand quickly up]*** up and ***[flips hand so that thumb is pointing downward, moves it quickly down]*** accelerate down, or does it just do one or the other?
233 S1: It just accelerates down.
234 T: Why do you say, just down?
235 S1: Because when it's going up, it's slowing down.

236 T: Mmm.
237 S1: Because it's getting ready to come back down the whole time, it's all about (*loudly*) coming down! It's never about the journey up!

Although the student was speaking loudly here, she appeared impatient rather than excited. When she repeated that the projectile was accelerating down, the teacher suggested that she ask herself whether that answer made sense to her. Her tone as she answered sounded impatient, if not flippant:

256 S1: That makes *so* much sense to me, we just *discussed* it.

The teacher left the table. Interestingly, no more than three seconds after the above comment, S2 pointed to a question on the activity sheet that asked about the direction of acceleration and asked her partner, apparently in all seriousness, “Is this up? Down-” (Figure 33).

In this small group discussion, in every instance where a student was observed

Figure 33: Persistent misconception about acceleration
Student on the right, pointing to her partner’s worksheet, “Is (*the acceleration*) up?”

exhibiting a conceptual difficulty, there was at least a brief response from either the other student or the teacher. Total discussion time spent on conceptual difficulties was ten and a half minutes, or 55% of the discussion time, most of which occurred when the teacher was present.

iii. Comparison: Year Two CP Teacher A

The videotape code maps (Figure 34) and Table 45 represent only Period 2 of the lesson sequence, during which the three Projectile Animations were used. In the code maps, the transcript segments run chronologically from left to right, spanning the time when the students were working with the animations and animation worksheets. Color blocks below each transcript segment denote the codes assigned to that segment; the codes are listed on the left. In these videotapes, the camera was used as a proxy for an individual student; the codes can be considered to reflect what an individual student in that class might have experienced.

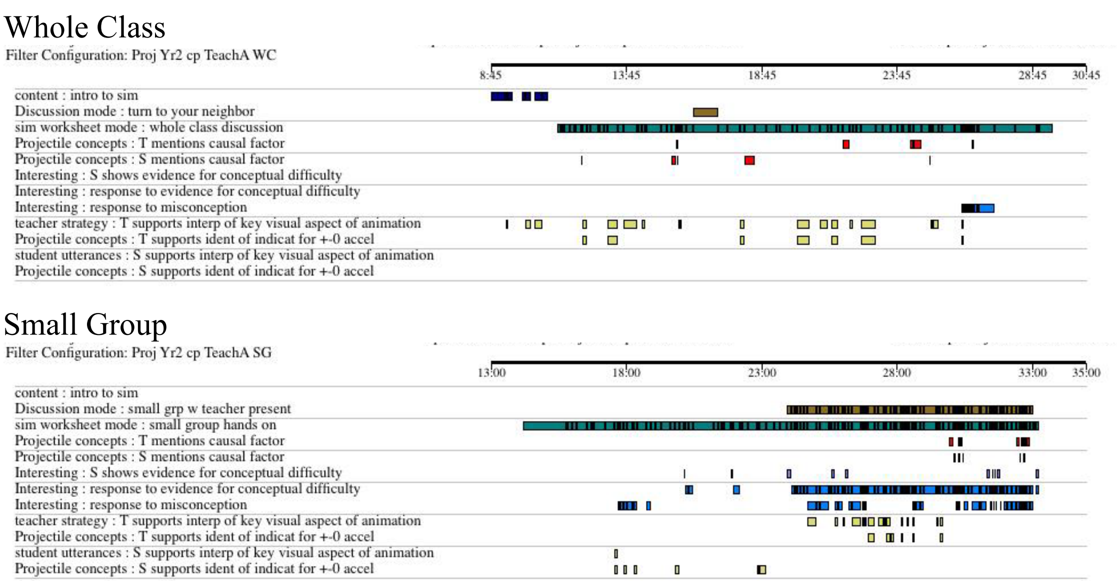


Figure 34: Videotape code maps: Year Two CP Teacher A
(Each timeline spans 22 minutes of videotape, not all of which was taken up by classroom discussion.)

Table 45: Videotape coding results: Year Two CP Teacher A

	Whole Class Format	Small Group Format
Time provided for activity sheets (Hand out until pick up)	24 min	23 min
Time provided for animations (including intro)	21 min	19 min + 1 ½ min whole class intro w/o activity sheets
Time utilized by students on camera for activity sheet questions (Starting at Q1)	18 min	19 min
Technical & other difficulties	0	0
Length of taped discussion analyzed below	18 min 21 sec	19 min 5 sec
Research Q #2: Discussion about key concepts	Total length: **1 min 30 sec** Percentage of discussion: **8%**	Total length: **57 sec** Percentage of discussion: **5%**
Research Q #3: Response to conceptual difficulties and misconceptions	Episodes of difficulty: 0 Response length: 0 Response to misc w no prior evidence of diff: Length: 1 min 16 sec Total: **1 min 16 sec** Percentage of discussion: **7%**	Episodes of difficulty: 10 Response length: 9 min 41 s Response to misc w no prior evidence of diff: Length: 54 sec Total: **10 min 35 sec** (Before T came: 2 episodes Response: 16 s, 16 s Response to misc: 54 s) Percentage of discussion: **55%**
Research Q #4: Support for key visual features	Total support episodes: **15** Teacher: 15 Student: 0 Avg: **49 per hour**	Total support episodes: **23** Teacher: 17 Student: 6 Avg: **72 per hour**

Discussion. The whole class and small group discussions observed for the College Prep (general level) physics classes appeared well matched across several of the parameters. They spent about the same amount of time on task and the teacher engaged in a similar number of visual support episodes in each discussion. Although the whole class spent more time and a greater percentage of time discussing the key concepts, neither discussion spent much time on these. In fact, in this matched class comparison, the greatest differences in results from video analysis were in favor of the small group.

Compared to the whole class discussion, **the small group on camera had**:

- more *student* visual support episodes (6 vs. 0);
- a greater *rate* of visual support episodes overall (student + teacher);
- 8x the *percentage* of discussion time addressing conceptual difficulties (also 8x the total time spent on difficulties).

However, the most notable feature of this small group was that

- **the teacher was present with the small group on camera for 47%** of the discussion time!

The small group visual support episodes, discussion of key concepts, and discussion about conceptual difficulties were all concentrated during the time the teacher was present. In fact, the teacher employed almost the same number of visual support moves during his 9 minutes with the small group as he had during the entire 21 minutes of the matched whole class discussion.

During the time when the teacher was not present, the results for this small group were consistent with those observed for many of the other small group discussions in this study:

- No discussion of key concepts;
- Low frequency of visual support episodes.

The total response to conceptual difficulties in the small group when the teacher was not present was short (less than a minute and a half over three episodes), but no shorter than in the matched whole class discussion (a minute and a quarter in a single episode).

Of course, the fact that the teacher spent so much time with the small group on camera meant that the four other small groups in the class, each of which had three students, shared the remaining 53% of the teacher's time between them. An estimate of how well these other students did can be obtained from the pre-post tests and activity sheets.

Considering all of the students in the classes, there was no significant difference between the pre-post gains in the two class formats for the short answer questions and the effect size was negligible [$t(21) = 0.34$, $p = 0.74$, $d = 0.14$]. (There was also no significant difference between the pre-test scores for the two groups.) Both conditions had significant gains with relatively large effect sizes ($d = 1.29$ and 0.89 respectively). Therefore, even though the teacher spent a lot of time with one small group and much less time with the other groups, the students in the small group class as a whole had very similar gains on the short answer questions to the students who had participated in the whole class discussion.

The gains for the discussion questions on the pre-post test appeared to follow a trend similar to that in this teacher's other matched sets; although no statistical analysis was done, the whole class students appeared to do better on this kind of question. They showed average gains of 11% whereas the students in the small group class actually showed an average *loss* of -6%.

Although there appeared to be more support for key visual features in the small group discussion on camera than in the whole class discussion, it is unknown what kinds of visual support occurred in the other small groups. However, the performance on the activity sheet questions, which addressed the meanings of the key features, can give an estimate of how much students were actually able to *use* the features. This performance was similar between the whole class and small group students, with the whole class averaging 64% correct and the small group students averaging 59%.

As with this teacher's other classes, the small group and whole class conditions appeared to have different strengths and weaknesses that largely balanced out; the evidence does not support the existence of any over-all advantage for the small group condition.

e. Year Two: Advanced Placement Physics (Teacher B)

Teacher B taught the Projectile Motion lesson sequence to the AP classes as a one-period lesson. The Galileo Simulation and the Projectile Animations were shown on the same day, with the three animations shown first. The part of the discussion that accompanied work with the animations and animations activity sheet was coded and analyzed. Although these students had not had the prior benefit of experience with the Galileo Simulation at the time they discussed the animations, this was equally true of both the whole class and small group conditions. Moreover, these students tended to be more advanced in their physics knowledge than the Honors class students (and considerably more so than the College Prep students) and, in general, appeared more ready to interpret and reason with the animations. The remainder of each class, which concerned the interactive Galileo Simulation, was transcribed and read to see whether any student frustration expressed and ignored during the animations discussion was addressed at any point before the post-test at the end of the period.

i. Whole Class Condition: Year Two AP Teacher B

The teacher introduced the lesson by throwing balls around the room and asking students, "What's happening? What are they doing? What's their motion?" The students and teacher mentioned causal factors such as gravity, the launching force, and air resistance. The teacher demonstrated the Galileo Simulation, which would be used during the second half hour of the class, and introduced unfamiliar terms. The students filled out prediction sheets. The teacher then turned to the Projectile Animations, giving a 2-minute introduction in which she used analogies and gestures to orient the students to the visuals on the display screen. At that point, she opened into a whole class discussion about the animations and animations activity sheet.

Shortly after the mid-point of the animations discussion, the teacher suggested to the class that they turn to their neighbors as they fill out their activity sheets. This initiated a 4-minute period where most students worked alone on their sheets, although a few occasionally spoke quietly with their neighbors. The remaining time in the animations discussion was an active 6-minute block that constituted the heart of the discussion and a short wrap-up discussion when the students discussed the answers they had written. Length of whole class discussion accompanying work with the animations and animations activity sheet was about 13 minutes. Analysis is in terms of the Research Questions 2, 3, and 4.

Research Question #2: Discussion about key concepts: Identifying concrete causal factors. Concrete causes were discussed rather extensively during the introduction to the lesson. During work with the Projectile Animations activity sheets, such causes were again mentioned but only briefly, twice by students and once by the teacher.

In the first episode, the teacher asked what was going on with the horizontal lines in Lines Animation I. A student replied that the distance between the lines was getting smaller because the velocity of the projectile was decreasing, because gravity was acting on it. In the second episode, students were trying to identify what, in Lines Animation I, let them know that the direction of acceleration was vertical and negative, and the teacher in passing mentioned gravity as a cause she knew they were familiar with. Immediately after this, in the third episode, the teacher asked the class whether acceleration was occurring in the horizontal direction and several students replied no. One student explained that this was because no forces were acting in that direction. This last episode was counted as mention of a causal factor even though it was actually the *lack* of a concrete cause resulting in the *lack* of an effect.

Total time spent discussing concrete causal factors during the whole class discussion was less than a minute. However, the introductory discussion about causal factors while the teacher and students were tossing balls around the room may have been sufficient for this Advanced Placement class.

Research Question #3: Response to conceptual difficulties and misconceptions. There was no evidence for conceptual difficulty observed during this discussion. This could have been because the advanced placement students did not want to bring up their difficulties in front of the class; however, the discussion as a whole appeared to go smoothly, with most students appearing to understand the concepts.

Research Question #4: Support for key visual features: Lines, arrows, spaces, and acceleration indicators. There was extensive support for the interpretation of visual elements in the animation, with one student and 18 teacher episodes observed. An extended example follows, involving both the student episode and three of the teacher episodes.

The student support move occurred at the very beginning of the discussion, when the Vectors Animation had just begun playing. The teacher opened the discussion, asking about the animated arrows that represented the vector components of the projectile's velocity. (Square brackets indicate gestures and boldface indicates depictive gestures. Underlining indicates utterances and gestures coded as visual support moves.)

47 T: It (*the worksheet*) is asking you what do you think the arrows indicate about velocity. … (W)hat do you think each of those arrows is representing?

48 *(Several students raise hands.)*

49 S: I think that the vertical one is the velocity for going up and down. So that ***[with left hand, points toward screen and traces an upward arc with her forefinger]*** as you're going up, the ***[suddenly flips her hand so that her forefinger is pointing downward]*** acceleration is ***[moves forefinger downward]*** negative, so the arrow is getting smaller. And then, like the acceleration of gravity, once you hit the peak, ***[moves forefinger slightly up and then back downward]*** the velocity is negative, and the ***[moves forefinger downward]*** acceleration of gravity is negative, too.

In her supporting move, the student used gestures and words to try to convey her understanding to the rest of the class, *selectively pointing out relationships to help others identify them as an indicator of acceleration* ("acceleration is negative, so the arrow is getting smaller") and *gesturing in the air to indicate the relationships*. This was an impressive move, a rare instance where a student identified visual relationships in the animation that indicated the presence, the direction, *and* the sign of the acceleration; it was negative both before and after the apex. About the only bit of information possible to derive from the vertical arrow this student did not mention was that it was possible to tell that the acceleration was *constant* and downward because the movement of the tip of the arrow relative to its tail was *constant* and downward. (No student gave evidence for noticing this relationship in any of the observed discussions.)

An interesting question might be where the teacher could go from here. The student had essentially given the answers. How could the teacher provoke reasoning on the part of students in the class who had not yet mentally engaged with the animation?

Rather than keeping the focus of the discussion on the direction of acceleration, a difficult concept for many students, the teacher returned the focus to the visual appearance of the vertical arrow and asked students what it might represent in terms of components.

50 T: So that's the one that they're calling Arrow A? So if we were to talk about that arrow, it gets smallerrrr *(looks toward display screen, animation apparently still running on loop)* and then it gets biggerrrrr and it's got something to do with the acceleration- What would you call that? As a component?

51 S: "j hat."

52 *(T laughs. Student has referred to the notation for unit vectors, $\hat{i}$, $\hat{j}$, $\hat{k}$.)*

53 T: It's the "j hat," in the "j hat" direction, right? But what is it? Is it position?

54 S: Velocity.

55 T: It's sort of like the velocity component? Maybe? Yeah? What about the horizontal one, the other one, Arrow B? What's that one doing?

56 S: Staying the same.

57 T: And what does it represent?

58 S: The horizontal velocity.

59 T: So if I wanted to, say, if I wanted to freeze this, right- unh! -there! (*pauses animation*) and I said, OK, so you're saying, horizontal component [*holds sheaf of papers up against the display screen immediately below the image of the projectile and its velocity vectors*], vertical component of velocity. What's the actual velocity at this moment in time?

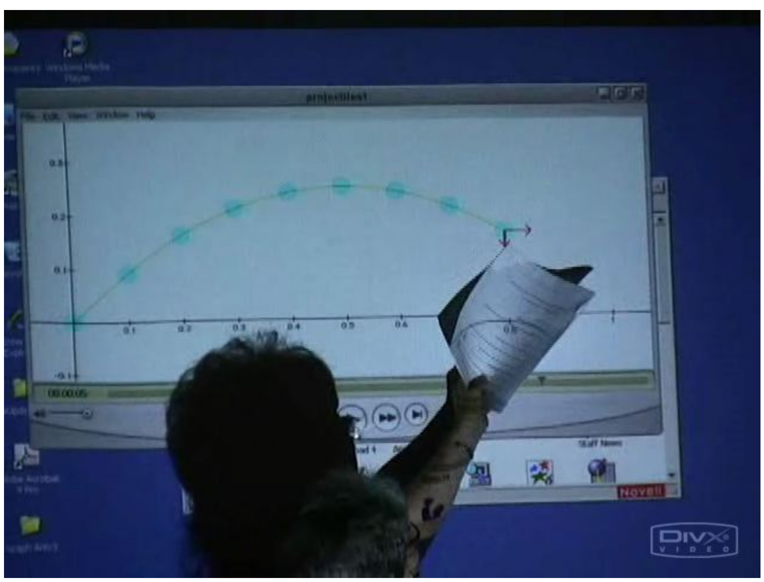


Figure 35: Teacher B uses a sheaf of papers to point out a key feature.

Here the teacher has provided several visual support moves. In Line 50, she *asked a question to prompt students to notice the change in size of Arrow A and to interpret its meaning* as a component. In Line 53, she *asked a probing question to prompt students to interpret the visual feature* not just in terms of mathematical formalism but also in terms of the physical aspect that it represented. In Line 59, she *selectively pointed out a relationship between key features* (Arrow A and Arrow B) at a specific point in time, by pausing the animation and by holding a sheaf of papers against the display to highlight those features. (See Figure 35.) She also *gave a hint to encourage the interpretation of the meaning of the arrows* by asking what was the actual velocity at that point in time. (Answer: the actual velocity is the resultant of vector addition of the two arrows.) The two moves in Line 59 were counted as a single support episode.

There were 19 support episodes during this whole class discussion, for an average rate of 95 per hour.

ii. Small Group Condition: Year Two AP Teacher B

The teacher began as in the whole class condition by throwing balls around the room and asking students what was influencing the motion. They mentioned causal factors such as gravity, the launching force, and air resistance. The teacher demonstrated the Galileo Simulation, which the class would be using during the second half hour of the period, and introduced unfamiliar terms. The students filled out prediction sheets. The teacher then introduced the Projectile Animations by describing them but did not bring them up onto the screen. Instead, she used analogies and gestures to describe what the students would see. The students then moved into their small groups.

In the small group on camera, S3 attempted to act as an "authority," though the other students did not always defer to him. Rather than attempting to support his group mates in gaining understanding, S3 tended to make factual pronouncements, not all of them correct. After a few seconds of confusion about which animation to bring up first, the students turned to the animations activity sheet and spent about 16 minutes in small group discussion on this activity sheet before turning to the second half of the lesson.

Research Question #2: Discussion about key concepts: Identifying concrete causal factors. During the animations portion of the small group discussion, concrete causal factors were mentioned in three episodes for a total time of about 20 seconds, or 2% of the discussion time. As with the matched whole class discussion, this may have been due to the fact that concrete causes had already been discussed rather extensively during the introduction to the lesson. However, it was not clear that all three of these students were comfortably able to distinguish which causal factors are in play during projectile motion.

The three episodes occurred close together; the first two are in Lines 105 and 106 below. S3 had already replied that there was downward acceleration being shown in the Vectors Animation (Video Clip 1), but the worksheet asked, "How do you know?" meaning, "What in the animation lets you know there is downward acceleration?" There were several possible answers in terms of visual representations in the animation. However, the discussion turned to underlying causal factors before the students had developed a clear understanding of the meaning of elements in the visual representation before them.

In the transcript excerpt, square brackets indicate gestures and boldface indicates depictive gestures. The excerpt begins as S2 read the question from the worksheet.

102 S2: "How do you know?" *(Question means, "How do you know that there is downward acceleration?")*
103 S3: How do you know? Gravity-
104 S1: Because, if you see this (*getting up and leaning toward display*), it doesn't change. It doesn't [*points to display screen*] get to the ground faster than it ***[points up]*** got up to the- (*pause, looks at companions, grins*) I mean, yeah, I guess it-
105 S3: I mean, what else is acting on it though? The point is that it's falling so there is gravity.
106 S1: [*pointing to the beginning of the trajectory on the display*] There is acceleration at the beginning [*points to the end of the trajectory*] and then at the end. Because of the force.

In Line 104, S1 attempted to give an answer to the worksheet question in terms of the visuals on the screen as the question intended, but became stumped. In Line 105, S3 invoked gravity as a causal factor rather than trying to answer in terms of the visuals. In line 106, S1 tried to reason in terms of the forces acting on the projectile at the beginning and end of its trajectory. These forces are causal factors but they do not operate while the projectile is actually in flight. (What was happening at the beginning and ending of the trajectory appeared to be a confounding factor for some of the students in these classes, as it had been for students in the preliminary interviews.) A few moments later, in a third episode, S3 invoked gravity again, not explaining but just saying, "I'm just gonna say, 'it is gravity.'" This series of episodes continues below, discussed in terms of the conceptual difficulties exhibited by the students. S1's attempts at visual support will be returned to in the section after that.

Research Question #3: Response to conceptual difficulties and misconceptions. S1 and S2 expressed conceptual difficulties on several occasions. Although S3 was incorrect at times, he never expressed this as a difficulty. In the excerpt in the previous section, S1 appeared to exhibit several misconceptions and S3 attempted to respond by stating a concrete cause. The following continues where the previous excerpt left off, and follows the students through a series of difficulties, responses, and notable absences of response. The focus of discussion continued to be the question on the worksheet asking how one could tell whether or not there was acceleration in the Vectors Animation. As the excerpt continues, S3 responded to S1's comments about forces and acceleration at the beginning and ending of the trajectory.

107 S3: I uh, I don't think we're supposed to read into it that hard.
108 S2: Well, isn't that like-
109 S3: *(inaudible)* the animation get the *(inaudible).* Changing like *y*-component. That's enough for me. I mean, like, nothing else is making it-

Beginning with Line 105 (in the section above), the discussion focused on S1's misconceptions for 32 seconds. Next, S2 exhibited difficulty.

110 S2: Is this a position graph? Doesn't a slope like that mean it's accelerating? It's accelerating.
111 S3: If th- yeah.
112 S2: There is acceleration going up and then when it goes down gravity- well, gravity- I don't know.
113 S3: I don't know, I don't think we're supposed to read that hard about it. Like, there're two other graphs, that we're just supposed to get the idea done, so I'm just gonna say, "it is gravity."

In Line 110, S2 appeared to be reasoning correctly about a visual aspect of the animation that indicated acceleration, namely, the changing slope of the graph, but in Line 112, he appeared confused when trying to relate this to gravity. S3's response was to advocate writing down "there is gravity" and getting it done with. In trying to reason about what they were "supposed" to do, S3 appeared to have missed the point of the question. S2, however, clearly *did* picked up on the point of the question and re-read it emphatically,

114 S2: "How do you *know*?"

S3 gave a short, inaudible response. The students wrote quietly on their activity sheets. A moment later, it became clear that S2 was still concerned with the issue:

117 S2: I said, the parabolic shape shows that it is accelerating.
118 S3: *(in an authoritative tone)* A parabolic shape indicates change in slope, which is characteristic of some kind of acceleration.

The students resumed writing on their activity sheets. The discussion response to S2's expression of difficulty was 36 seconds.

A little later, S1 expressed confusion about a similar issue with respect to Lines Animation I, which had horizontal lines indicating the changing motion along the *y*-axis. The activity sheet question was, "What does the spacing between the lines indicate about the acceleration?"

155 S1: Should we just say, like, it decreases as it goes up, increases as it goes down? Like that?
156 S2: Huh?
157 S1: Should we do it like that? Like, it incr- it decreases as it goes up, and increases as it goes down? Just like that?
158 *(Pause as the students looked at the display screen, the animations still running on loop.)*
159 S1: It changes uniformly?

To this, his classmates gave no response. Instead, S3 made a comment about how the two animations they had up and running were starting to "desync," to run out of synchrony with each other (synchrony was not required for their interpretation). Then the students resumed writing on their activity sheets without talking.

A few moments later, S1 tried again to ask what was indicated by the spacing between the lines in the animation. S2 and S3 had apparently gone on to the questions for Lines Animation II, while S1 was still working on the questions for Lines Animation I. This resulted in confusion and miscommunication. S1 asked his question several more times, finally beginning to appear frustrated at non-responses such as the following.

171 S1: *(re-reading the question)* “What does the spacing- ”
172 S3: Magical dilemma of life, whether we know too much or not.
173 S1: *(sounding frustrated)* No, you can't just say something that (*inaudible*), make it sound like- (*inaudible*).

However, he appeared to rein in his frustration quickly and, with quiet persistence, continued to bring up the topic, asking repeatedly whether it wasn’t the acceleration that was decreasing rather than the velocity. After about 4 minutes, he finally dropped the issue after having raised it nine times. The students left the animations and moved on to the Galileo Simulation and the simulation activity sheet. Total response to S1’s repeated expressions of difficulty over the four-minute period was 17 seconds.

During the Galileo Simulation part of the discussion, which was transcribed but not coded, the conversation never returned to S1’s questions about acceleration and velocity. This is not surprising because the simulation explored a different aspect of projectile motion than did the Projectile Animations. However, there was an interesting resolution to the uncomfortable dynamic reflected in the transcript excerpts above. At one point S3 left the group. While he was away, S1 and S2 used the Galileo Simulation to discover a surprising fact about projectiles (unequal masses shot at equal velocities will travel the same path in the absence of air resistance) and S1 figured out a concrete cause for this phenomenon (one has to use more force with the more massive projectile to get it up to speed). When S3 returned, S1 and S2 asked him what his prediction on the prediction sheet had been for the case of unequal masses. When he answered incorrectly, S1 and S2 looked at each other, shrugged and laughed. S1 responded (with a note of glee?), “You will be pleasantly surprised.” After this, the dynamic of the group changed, S1 sounding more confident and not hesitating to point out when his predictions were correct.

Total time spent addressing student conceptual difficulties and misconceptions was about a minute and a half, or 9% of the discussion time. Other classes had a percentage this small or smaller, especially if there had been only one or two episodes of difficulty. A notable feature of this discussion, however, was that 10 episodes of expressed difficulty had produced such a small amount of response.

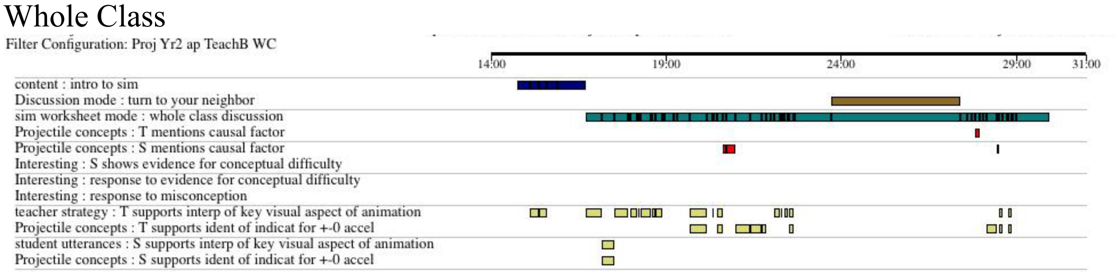
Research Question #4: Support for key visual features: Lines, arrows, spaces, and acceleration indicators. There was no support observed for the identification or interpretation of key visual features during the animations portion of this small group discussion. In Lines 104 and 106 above, S1 appeared to be attempting some kind of visual support but, although his references were not clear (perhaps not even to himself), he did not appear to be referring to any of the key features or relationships, but rather to some symmetry between the beginning and ending of the trajectory. The teacher had engaged in a brief series of visual support moves during her introduction, using analogy and gestures to explain to the students what they would be seeing in the animations. This was before the animations were brought up and before the students broke up into groups (and was not counted). Much later, she also gave some support to the small group on camera concerning visual elements in the Galileo Simulation (not counted in these comparisons). This lack of support for key visual features in the animations is especially interesting in view of the 'expert' role taken on by one of the students, and in view of the difficulties experienced by the other two students when attempting to use the visual features of the animations to reason about the acceleration of the projectile.

Other comments. An additional question is to what extent socio-cultural factors might have played a role in the performance of this small group; such an analysis, however, falls beyond the scope of the present study.

iii. Comparison: Year Two AP Teacher B

The videotape code maps (Figure 36) and Table 46 represent only the portion of this one-period lesson during which the three Projectile Animations were used. In the code maps, the transcript segments run chronologically from left to right, spanning the time when the students were working with the animations and animation worksheets. Color blocks below each transcript segment denote the codes assigned to that segment; the codes are listed on the left. In these videotapes, the camera was used as a proxy for an individual student; the codes can be considered to reflect what an individual student in that class might have experienced.

Whole Class



Small Group

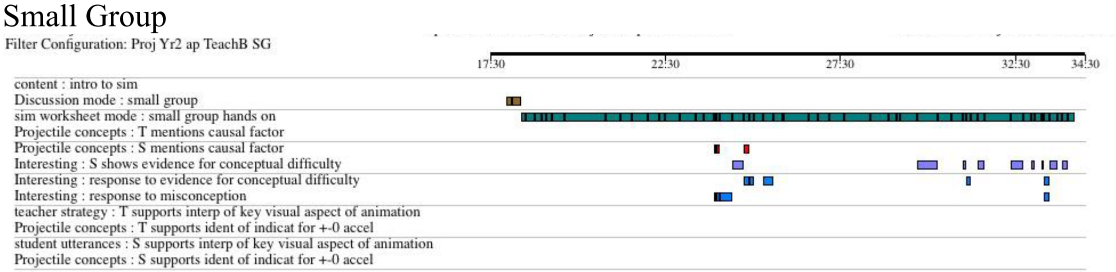


Figure 36: Videotape code maps: Year Two AP Teacher B
(Each timeline spans 17 minutes of videotape, not all of which was taken up by classroom discussion.)

Table 46: Videotape coding results: Year Two AP Teacher B

	Whole Class Format	Small Group Format
Time provided for activity sheets (Hand out until pick up)	15 min	19 min
Time provided for animations (including intro)	15 min	16 min
Time utilized by students on camera for activity sheet questions (Starting at Q1)	13 min 15 sec	16 min
Technical & other difficulties	0	39 sec
Length of taped discussion analyzed below	13 min 15 sec	15 min 49 sec
Research Q #2: Discussion about key concepts	Total length: **32 sec** Percentage of discussion: **4%**	Total length: **20 sec** Percentage of discussion: **2%**
Research Q #3: Response to conceptual difficulties and misconceptions	Episodes of difficulty: 0 Response length: 0 Response to misc w no prior evidence of diff: 0 Response length: 0 Total: 0 Percentage of discussion: 0%	Episodes of difficulty: 10 Response length: 53 sec Response to misc w no prior evidence of diff: 1 Response length: 32 sec Total: 1 min 25 sec Percentage of discussion: 9%
Research Q #4: Support for key visual features	Total support episodes: 21 Teacher: 20 Student: 1 Avg: 95 per hour	Total support episodes: 0 Teacher: 0 Student: 0 Avg: 0 per hour

Discussion. It can be seen from the above that the whole class had:

- **many visual support episodes** for key features **(**compared to none for the small group);
- no episodes of student difficulty (compared to **many for the small group**);
- not much discussion of key concepts (nor did the small group).

The visual contrast in these videotape code maps is especially striking. In the whole class discussion, the visual support episodes (gold color blocks) occur thickly except when the students were engaged in an extended "turn to your neighbor" session (actually "turn to your neighbor or write on your own") in which most students were writing on their activity sheets. In the small group discussion, many of the student conceptual difficulties (lilac color blocks) were not responded to (not followed by cyan color blocks), at least in any way that was visible or audible to the observer.

Neither class had much discussion about concrete causal factors once the animations were brought up. The classes seemed comfortable with gravity as an important causal factor in projectile motion (although not necessarily comfortable with it as *the* important causal factor), and this, together with the fact that causal factors were discussed fairly extensively during the introductions, may have been a reason they were not mentioned at any length during the observed discussions. However, when one student in the small group wanted to discuss other concrete causes that were not relevant to the discussion at hand, his request was not responded to in a meaningful way; a discussion about causal factors at that point might have been helpful.

Teacher B had her Advanced Placement classes reverse the order in which they encountered the Galileo Simulation and Projectile Animations, viewing the animations first. She also had them complete both activity sheets during a single class period. This length of time seemed about right for these two AP classes. The time on task for the whole class discussion was actually slightly less than the time on task for the small group on camera; this is the only comparison for which this was true. This was partly because the time on task was calculated from the beginning of discussion about Question 1 on the activity sheet, and in the whole class setting, the teacher delivered about two minutes of additional introduction to the Vectors Simulation after she brought it up before turning to the first activity sheet question. Total time spent with the animations was very similar in the two classes. (Beginning the timing at Question 1 provided the best consistency across the study because the teachers varied how and when they conducted their introductions to the visual tools, although this meant that visual support episodes occurring during the introductions were omitted from the comparisons.)

It appears clear that there were many more visual support episodes for key features in the whole class discussion than in the small group on camera. However, the small group on camera may not have been typical. An estimate of the extent to which students were able to *use* the key features and relationships in their own thinking can be obtained from the relevant questions on the activity sheet. On these, the whole class students averaged 65% correct while the small group students averaged 77% correct. Although no statistical analysis was done, the indication is that students in the small group condition as a whole appeared able to reason with the features at least as well as the participants in the whole class discussion and possibly better.

The small group on camera also may have been atypical in the amount of conceptual difficulty exhibited. Alternatively, it may be that these Advanced Placement students were more hesitant to bring up their conceptual difficulties in front of the whole class than they were in small group. Whatever the strengths and weaknesses for each format, the effect on pre-post gains appeared to balance out across the two classes. For the short answer questions, statistical analysis revealed no significant difference between the gains in the two classes [$t(39) = 0.04$, $p = 0.97$, $d = 0.01$]. Both groups had significant gains at the $p < 0.001$ level with relatively large effect sizes ($d = 0.89$ and 1.01 for WC and SG respectively) and each class attained approximately 65% of the gains possible for them, given their fairly high pre-test scores. Though no statistical analysis was done on the explanation questions, the gains of the two groups on these also appeared similar, at 19% for the whole class condition and 17% for the small group condition. Once again, although the strengths and weaknesses within the two formats appear different, there is no evidence for any overall pre-post advantage for the students in the small group format.

D. Global Patterns in Case Studies

The above analyses discuss the coding results within the context of each case study comparison. Although a statistical analysis across matched sets was not attempted, it is worth asking whether any global patterns in the coding can be observed. One apparent pattern is that the videotape code maps of the small group discussions appear to vary more than those of the whole class discussions. For instance, the length of discussion time varied especially widely across the small groups. The quality of discussion also appeared to vary widely, as can be seen from the case study descriptions:

- 4 small groups appeared to be well functioning;
- 3 groups had students who appeared either not to be interested or to be in "data collection mode," not interested in engaging in probing discussion;
- 2 groups appeared to have socio-cultural factors affecting the interactions in unhelpful ways, most notably the group in which the repeated questions of a member were ignored.

Another global pattern that can be seen in the videotape code maps is that the codes for visual support episodes and for discussion about key concepts appear most frequently along with the codes for the presence of a teacher. For the whole class discussions, these codes tend to be distributed throughout the discussions, while for many of the small group discussions, the codes are clustered around the teacher visits to the group. In some cases, even the student support episodes are clustered around the teachers' visits, as with the group on camera in Teacher C's class. Several possible explanations for this will be discussed in Chapter VIII.

CHAPTER VIII

CONCLUSIONS

An important factor that cut across the five research questions was that general assumptions about the advantage of small group, hands on work at computers over whole class use of computers have not been tested. Therefore, each of the questions was investigated in the context of both small group and whole class use, and assumptions about the advantage of one lesson format over the other were examined in light of the results. Results that can address each research question are collected and presented here, followed by a bulleted list of findings and discussion.

A. Research Question #1

Are students able to exhibit gains in conceptual reasoning on pre-post tests from lessons that incorporate certain interactive simulations and animations?

Eighteen of nineteen classes in the study had statistically significant pre-post gains on transfer short answer questions that addressed conceptual issues, $p < 0.025$, effect sizes ranging from $d = 0.42$ (small) to $d = 1.77$ (large). The class that did not have significant gains was an Advanced Placement small group class that had fairly high scores on the pre-test at 76% and gains of 2% ($p = 0.44$, $d = 0.18$). This raises the question of whether the lesson was appropriate for this group. However, the matched whole class discussion condition also had pre-test scores of 76% and had 10% gains significant at $p < 0.001$ with a medium effect size, $d = 0.60$, suggesting that there was sufficient room in these classes for gains.

Statistical analysis was not conducted on the scoring of the pre-post discussion questions because there were only one or two per test. These were scored blind to condition according to evidence for presence of certain key concepts. Some but not all classes had gains on the discussion questions, ranging from -6% to 43%.

Examining these data in the context of whole class (WC) and small group (SG) discussion conditions, a pattern begins to emerge. The results of statistical comparisons of short answer gains are summarized in Tables 47 and 48, while gains on explanation questions are listed without statistical analysis in Tables 49 and 50.

Table 47: Gravitational PE short answer transfer question pre-post gains

	WC Gains			SG Gains						
	N	Mean	SD	N	Mean	SD	*t*-Value	Sig.	Cohen's *d*	
HP	20	**0.22**	0.21	19	0.09	0.15	2.221	0.033*	0.71	**
CP	11	**0.26**	0.20	14	0.25	0.24	0.097	0.924	0.04	
AP	23	**0.10**	0.12	21	0.02	0.11	2.368	0.023*	0.71	
AP	21	**0.09**	0.16	21	0.07	0.10	0.506	0.616	0.16	

Boldface indicates the larger mean gain within each matched set.
*Significant difference in gains in favor of the whole class condition.
**Unanticipated events may have had a disproportionate effect on the small group condition; those students were encouraged to finish their activity sheets in a single period while students in the whole class condition were not.

Table 48: Projectile Motion short answer transfer question pre-post gains

	WC Gains			SG Gains					
	N	Mean	SD	N	Mean	SD	*t*-value	Sig.	Cohen's *d*
HP	21	0.35	0.35	25	**0.36**	0.34	-0.087	0.931	0.03
HP	34	**0.35**	0.29	19	0.32	0.24	0.294	0.770	0.08
HP	15	**0.41**	0.22	22	0.37	0.33	0.471	0.640	0.16
CP	14	**0.31**	0.29	9	0.27	0.29	0.336	0.741	0.14
AP	20	0.22	0.22	21	0.22	0.23	0.036	0.971	0.01

Boldface indicates the larger mean gain within each matched set.

Ignoring the result in the shaded row, there is still no evidence in these results for a pre-post short answer advantage for the small group hands-on condition. If anything, there appears to be a slight trend in favor of the whole class condition, especially in the Gravitational Potential Energy classes. The effect sizes for most differences were negligible, but in one comparison, the difference in gains reached significance in favor of the WC condition with a medium effect size, with no obvious reason for the lower scores in the SG condition other than difference in condition. There was no ceiling effect here, for example. This SG class was the only class in the study that did not exhibit significant pre-post gains on the short answer questions.

Table 49: Gravitational PE explanation question pre-post gains

	WC Gains		SG Gains		
	N	Mean	N	Mean	
HP	20	0.14	19	**0.26**	*
CP	11	0.05	14	**0.21**	
AP	23	**0.24**	21	0.02	
AP	21	**0.43**	21	0.26	

Boldface indicates the larger mean gain within each matched set.
* Although unanticipated events may have had a disproportionate effect on the small group condition (see Table 47), no evidence for that is seen here.

Table 50: Projectile Motion explanation question pre-post gains

	WC Gains		SG Gains	
	N	Mean	N	Mean
HP	21	**0.19**	25	0.11
HP	34	**0.17**	19	0.12
HP	15	**0.17**	22	0.06
CP	14	**0.11**	9	-0.06
AP	20	**0.19**	21	0.17

Boldface indicates the larger mean gain within each matched set.

Although statistical analysis was not conducted on the gain scores for the explanation questions, a trend can be observed in the five Projectile Motion comparisons; in every Projectile Motion comparison, the whole class had larger mean gains on explanation questions. Although no such trend is observed for the Gravitational Potential Energy scores, it seems reasonable to conclude that there is no evidence here for any over all advantage conferred by the small group condition. These results reinforce the results for the short answer questions in Tables 47 and 48 above; in general, pre-post gains for students in the two conditions appear to be very close with a slight trend in favor of the whole class condition.

Interestingly, the HP small group class that was flagged in Table 47 due to a possible disproportionate affect on the students of a teacher action does *not* show smaller gains on the explanation questions. This is the kind of mixed result that led to a desire to conduct videotape analysis, the results of which are discussed next.

B. Research Question #2

To what extent do students and teachers engage in discussion about key concepts while working with the simulations and animations?

The key concepts identified in the preliminary studies were, for the Gravitational Potential Energy lesson, a) the possibility that the total energy of a system could equal zero and b) the possibility that some kind of energy value could take on negative values. For the Projectile Motion lesson, key concepts involved concrete causal explanations for aspects of projectile motion. These included a) the presence of gravity (a gravitational force, to be precise) as a concrete explanation to account for the presence of acceleration in the vertical direction, and b) the absence of outside forces in the horizontal direction as a concrete explanation for the constant velocity of the projectile in that direction. The amount of discussion and percentage of discussion time devoted to these concepts are given below.

Table 51: Gravitational PE discussion about key concepts
Expressed as percentage of discussion time

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 HP	Teacher A	2.85 min / 62.03 min = **0.05**	0.75 min / 29.23 min = 0.03
Yr 1 CP	Teacher B	4.32 min / 42.42 min = **0.10**	0.40 min / 23.90 min = 0.02
Yr 1 AP	Teacher B	0.92 min / 41.10 min = 0.02	0.99 min / 32.32 min = **0.03**
Yr 2 AP	Teacher B	2.58 min / 41.71 min = **0.06**	1.16 min / 28.95 min = 0.04

Results expressed in minutes, *not* in minutes and seconds. Boldface indicates the larger percentage in each matched set.

Table 52: Projectile Motion discussion about key concepts
Expressed as percentage of discussion time

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 HP	Teacher A	0.25 min / 15.88 min = 0.02	1.83 min / 12.02 min = **0.15**
Yr 1 HP	Teacher C	0.33 min / 15.62 min = **0.02** 2.97 min / 18.55 min = **0.16**	0.00 min / 17.23 min = 0.00
Yr 2 HP	Teacher A	0.62 min / 36.00 min = 0.02	0.65 min / 12.42 min = **0.05**
Yr 2 CP	Teacher A	1.50 min / 18.35 min = **0.08**	0.95 min / 19.08 min = 0.05
Yr 2 AP	Teacher B	0.53 min / 13.25 min = **0.04**	0.33 min / 15.82 min = 0.02

Results expressed in minutes, *not* in minutes and seconds. Boldface indicates the larger percentage in each matched set.

For both lesson sequences, the key concepts were discussed for at least a few seconds in every discussion with the exception of one of the small group discussions. However, in that class, the teacher led an extended wrap-up discussion in which the key concepts were discussed after the small group discussions were over. In general, though, there was little discussion time devoted to these concepts, ranging from 2% to 16% of the discussion time in the whole class discussions and from 0% to 15% in the small groups on camera.

In 6 of the 9 comparisons, the percentage of discussion time spent on the key concepts was greater in the whole class discussions than in the matched small group discussions. Considering that the whole class discussions tended to spend longer on task, the raw time spent on the key concepts was substantially greater in these whole class discussions. In two other comparisons the percentage of time was similar in the two conditions. In only one comparison did the small group on camera spend substantially more time discussing the key concepts than did the matched whole class discussion. However, all of these times were brief. The most time spent on discussing these concepts in any class was 4 1/3 minutes in a whole class Gravitational Energy discussion; in a majority of the classes, such discussion lasted less than a minute. The small size of these numbers was surprising, given that the simulations and animations provided important potential affordances for developing these key concepts.

The evidence described here does not suggest an advantage for the students in the small group condition; if anything, the trend was in favor of the whole class discussions.

C. Research Question #3

To what extent do teachers and students respond to conceptual difficulties and misconceptions exhibited during work with the simulations and animations?

The intention was to identify when student conceptual difficulties were being acknowledged and addressed by the discussion and to look for possible patterns in such support. The amount of time and percentage of discussion time spent addressing student difficulties is given in Tables 53 and 54 below.

Table 53: Gravitational PE response to conceptual difficulties
Expressed as percentage of discussion time

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 HP	Teacher A	14.05 min / 62.03 min = **0.23**	3.35 min / 29.23 min = 0.11
Yr 1 CP	Teacher B	6.15 min / 42.42 min = **0.14**	0.89 min / 23.90 min = 0.04
Yr 1 AP	Teacher B	3.72 min / 41.10 min = **0.09**	1.58 min / 32.32 min = 0.05
Yr 2 AP	Teacher B	1.79 min / 41.71 min = 0.04	3.12 min / 28.95 min = **0.11**

Results expressed in minutes, *not* in minutes and seconds. Boldface indicates the larger percentage in each matched set.

In three of the four comparisons, the whole class discussion spent considerably longer responding to student conceptual difficulties, all of them at least double the time spent in the matched small groups. Even accounting for the shorter time on task for these small group discussions, the percentage of discussion time spent addressing student difficulties was less. In one class, however, both the amount and percentage of time addressing difficulties was greater in the small group; this appeared to be an exemplary small group.

Table 54: Projectile Motion response to conceptual difficulties
Expressed as percentage of discussion time

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 HP	Teacher A	2.42 min / 15.88 min = **0.15**	1.35 min / 12.02 min = 0.11
Yr 1 HP	Teacher C	2.88 min / 15.62 min = 0.18 1.38 min / 18.5 5min = 0.07	2.85 min / 17.23 min = 0.17
Yr 2 HP	Teacher A	15.48 min / 36.00 min = **0.43**	4.77 min / 12.42 min = 0.38
Yr 2 CP	Teacher A	1.27 min / 18.35 min = 0.07	10.58 min / 19.08 min = **0.55**
Yr 2 AP	Teacher B	0.00 min / 13.25 min = 0.00	1.42 min / 15.82 min = **0.09**

Results expressed in minutes, *not* in minutes and seconds. Boldface indicates the larger percentage in each matched set.

In the Projectile Motion lessons there is no clear trend. In the AP whole class discussion, none of the discussion addressed student difficulties. (There were also no expressions of student difficulty observed in that discussion). In two other matched sets, there were large differences between the conditions in the amount of time spent on addressing difficulties. In one, it was the whole class discussion that spent more time, the one in which Teacher A had given a long, impromptu lecture-demonstration in response to difficulties. In the other matched set with a large difference, it was the small group discussion that spent more time addressing difficulties. This was the small group discussion in which Teacher A had spent half the discussion time with the group on camera helping the students address their conceptual difficulties.

In summary, the whole class discussions spent longer responding to student difficulties in five of the comparisons and the small group discussion spent longer in four of them (one of which had the teacher present). The evidence here does not suggest an advantage for the small group students in getting their conceptual difficulties addressed by discussion.

D. Research Question #4

To what extent do teachers and students support the recognition, use, and interpretation of key visual features of the simulations and animations?

The intention was to identify the amount and kinds of support used to address student *perceptual* and other difficulties in making effective use of key visual features (see Research Question 5 below for an explanation of these) that were intended affordances of the simulations and animations. Either a teacher or student could employ these support moves. Preliminary studies led to the identification of an initial set of support moves that appeared to have been useful for the students in those classes. The set of support moves and their descriptions were further developed and honed through an iterative process involving analysis of additional videotapes and transcripts during the main study. The honed list of visual support moves constitutes a finding of this study[11]:

1) **Selectively pointing out some aspect of the key visual feature or relationship** as part of an apparent attempt to help students use it or interpret its meaning**;**

2) **Giving a hint to encourage use or interpretation of the meaning** of the key visual feature or relationship;

3) **Gesturing in the air or over the display** to indicate the key visual feature or relationship as part of an apparent attempt to help students use it or interpret its meaning;

4) **Asking a question to prompt use or interpretation** of the meaning of the key visual feature or relationship;

5) **Suggesting a manipulation of the simulation** to assist with use or interpretation of the meaning of the key visual feature or relationship;

[11] Sample video excerpts illustrating these support moves may be available to educators; contact the author.

6) **Pointing out a limitation to interpreting the meaning** of the key visual feature or relationship.

The numbers and frequencies of visual support episodes are given in Tables 55 and 56 below. (Additional moves that were identified but that were rarely seen or were too context specific to be compared across lesson formats were catalogued and are a possible subject for future research. Some of these were listed in Chapter III, Section B.3.)

Table 55: Gravitational PE support for key visual features
Expressed as episodes per hour

Class	Teacher	Whole Class Format		Small Group Format	
Yr 1 HP	Teacher A	26 / 62.03 min =	**25** per hour	8 / 29.23 min =	16 per hour
Yr 1 CP	Teacher B	37 / 42.42 min =	**52** per hour	4 / 23.90 min =	10 per hour
Yr 1 AP	Teacher B	19 / 41.10 min =	**28** per hour *	10 / 32.32 min =	19 per hour *
Yr 2 AP	Teacher B	19 / 41.71 min =	**27** per hour	10 / 28.95 min =	21 per hour

Boldface indicates the larger rate within each matched set.
*The students in the small group engaged in support for other features that they adapted in place of key features they couldn't find. This number represents the amount of support for key features only, but under-represents the amount of visual support these small group students actually provided each other.

In no comparison did the small group discussion show an advantage. In the shaded row, with a different coding scheme, the small group might have been close to the whole class in numbers of episodes, but with no scheme trialed in this study would they have exceeded the whole class. In the other three comparisons, the whole classes clearly exceeded the matched small groups in rates of support episodes, and far exceeded them in raw numbers of episodes.

In the Projectile Motion lessons (Table 56 below), the pattern is not so clear. However, in three of five comparisons the whole class discussion had a greater rate of support episodes, and in four of five, they had a greater total number of episodes. Importantly, in the small group that had a greater total number of support episodes, *all* *but 6 of the 23 episodes were teacher episodes that occurred while the teacher was with the small group.*

Table 56: Projectile Motion support for key visual features
Expressed as episodes per hour

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 HP	Teacher A	12 / 15.88 min = 45 per hour	11 / 12.02 min = **55** per hour
Yr 1 HP	Teacher C	54 / 15.62 min = **207** per hour 40 / 18.55 min = **129** per hour	25 / 17.23 min = 87 per hour
Yr 2 HP	Teacher A	25 / 36.00 min = **42** per hour	1 / 12.42 min = 5 per hour
Yr 2 CP	Teacher A	15 / 18.35 min = 49 per hour	23 / 19.08 min = **72** per hour
Yr 2 AP	Teacher B	21 / 13.25 min = **95** per hour	0 / 15.82 min = 0 per hour

Boldface indicates the larger rate within each matched set.

Note that episodes of mutual support *were* counted; it was not required that the person engaging in support be acting as an expert, only that the move appeared intended to help other students in addition to the supporter. Even so, *the trend here seems to be in favor of discussion in which the teacher is present,* and, perhaps consequently, leans toward whole class discussion, although it is certainly possible that exemplary small groups may be able to support each other within group.

E. Research Question #5

Do students recognize and use key visual features of the simulations and animations?

Whether or not students were *supported* in their use of key features (Research Question 4), an important issue is whether they actually *recognized and used* them, although this question is more difficult to address. Preliminary studies led to the identification of an initial set of visual features that appeared useful to the students in those studies as they attempted to reason about the key concepts (described under Research Question 2). These included: the gravitational potential energy (GPE) reference line and the animated energy bar graph in the PhET Energy Skate Park Simulation, and the animated arrows, vertical and horizontal lines, and various visual relationships in the Projectile Animations.

Student answers to selected activity sheet questions were coded (blind to condition) for evidence for use of these key features and relationships. For the Gravitational Potential Energy lessons, where only one or two questions per sheet were coded, videotape coding allows for an alternate snapshot of student use; the results of the two kinds of analyses can be examined in light of each other (Tables 58-60). For Projectile Motion, no videotape coding was practical because the features could not be manipulated. However, the Projectile Activity Sheets were designed so that four questions directly asked for interpretation of the key features, and student answers to these questions could be scored from a simple rubric as correct, partially correct or incorrect. Because these results are simpler than those for Gravitational Potential Energy, they are presented first (Table 57).

Table 57: Class Mean Performance on Four Projectile Activity Sheet questions
Means expressed as percentages of a perfect score.

		WC		SG	
Class	Teacher	N	Mean	N	Mean
Yr 1 HP	Teacher A	21	0.63	25	**0.75**
Yr 1 HP	Teacher C	17 19	**0.82** **0.68**	19	0.66
Yr 2 HP	Teacher A	18	0.74	24	**0.75**
Yr 2 CP	Teacher A	14	**0.64**	10	0.59
Yr 2 AP	Teacher B	22	0.65	23	**0.77**

Boldface indicates the larger mean score within each matched set.

The class means for student performance on the four relevant Projectile Activity Sheet questions are given in Table 57 as percentages of a perfect score. Even though statistical analysis of these data is not attempted, it can be seen that the means for all classes were greater than 50%. This suggests that many of the students were able to recognize and use the visual features that were designed to give information about the presence and direction of acceleration in the system; however, no clear whole class/small group pattern emerges. In two of the matched sets, the whole class means are higher; in two others, the small group means are higher; and in a fifth matched set, the means of the two conditions are essentially the same. These data do not suggest an overall advantage for the small group students in being able to use the visual features of the animations in their thinking, which may not be surprising given the nature of the animations; they provided limited interactivity, often considered the forte of small group work.

The Gravitational Potential Energy results concern a situation where extensive manipulation of the features was possible. For this lesson sequence, videotape analysis for use of these features was employed, allowing for a snapshot of student use that can be considered along with the activity sheet results to be discussed below. Results for videotape coding for recognition and use of the features are in Tables 58 and 59.

Table 58: Gravitational PE student and teacher recognition and use of key visual features

Expressed as episodes per hour

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 HP	Teacher A	31 / 62.03 min = 30 per hour	21 / 29.23 min = **43 per hour**
Yr 1 CP	Teacher B	58 / 42.42 min = **82 per hour**	11 / 23.90 min = 28 per hour
Yr 1 AP	Teacher B	21 / 41.10 min = 31 per hour	18 / 32.32 min = **33 per hour**
Yr 2 AP	Teacher B	37 / 41.71 min = 53 per hour	34 / 28.95 min = **70 per hour**

Boldface indicates the larger rate within each matched set.

Videotape coding identified *more episodes* of recognition and use of key visual features in each whole class discussion than in the matched small group discussion, but this may have been due to the fact that the small groups spent shorter time on task. For the *rate* of use, the small groups on camera average a greater rate in two of the comparisons and about the same rate in another, although the whole class discussion had a far greater rate in the fourth comparison. Another factor is that much of the use in the whole class condition was by the teacher. However, although it was hoped to be able to compare *student* use in the two conditions, it was difficult to detect evidence of student recognition and use in the whole class condition, especially because the camera was in the back of the room during whole class discussions. Perhaps it is not surprisingly that when such a comparison is attempted, much more evidence for student use is visible in the videotapes of small group activity (Table 59).

Table 59: Gravitational PE student-only recognition and use of key visual features
Expressed as episodes per hour

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 HP	Teacher A	5 / 62.03 min = 5 per hour	21 / 29.23 min = **43 per hour**
Yr 1 CP	Teacher B	25 / 42.42 min = **35 per hour**	9 / 23.90 min = 23 per hour
Yr 1 AP	Teacher B	5 / 41.10 min = 7 per hour	13 / 32.32 min = **24 per hour**
Yr 2 AP	Teacher B	5 / 41.71 min = 7 per hour	33 / 28.95 min = **68 per hour**

Boldface indicates the larger rate within each matched set.

It is interesting to note that in the lower level College Preparatory (CP) class, students in the whole class discussion were observed using the key features *more times* and at a *greater rate* than the students in the matched small group on camera, even though, again, the camera was behind the students in the whole class discussion. However, the CP small group data is only for a single small group containing two students. Despite the CP data, the results presented in Tables 58 and 59 appear to go against the trend favoring the whole class condition otherwise suggested by this study. Perhaps even more interesting are the large group data as a whole: the large number and high rate of identification and use of key features among the CP students in the whole class condition were much higher than the numbers and rates in the whole class discussions of the more advanced HP and AP students.

An alternate snapshot of student use can be provided by activity sheet analysis, although this analysis considered only the one or two questions on each sheet that had the potential to provide evidence for use of the features (one question on the Year 2 AP activity sheet and two questions on the sheet used by all six of the Year 1 classes); this analysis has the advantages of cutting across all of the students in these classes and that it could be conducted blind to condition. Note that Table 60 is laid out with whole class data listed above the matched small group data rather than beside it.

Table 60: Class Mean Performance on Gravitational PE Activity Sheet Questions 6 and 7

Means expressed as percentages of students who showed the evidence described.

Class	Teacher	N	Lesson Format	1) evidence for use of GPE reference line	2) evidence for use of concepts supported by bar graph	3) evidence for use of key relationship	
Yr 1 HP	Teacher A	20	WC	**0.10**	**0.05**	**0.05**	
Yr 1 HP	Teacher A	18	SG	0.00	0.00	0.00	
Yr 1 CP	Teacher B	11	WC	**0.36**	**0.27**	**0.18**	
Yr 1 CP	Teacher B	13	SG	0.00	0.00	0.00	
Yr 1 AP	Teacher B	13	WC	0.15	0.23	0.08	*
Yr 1 AP	Teacher B	18	SG	**0.33**	**0.44**	**0.22**	
Yr 2 AP	Teacher B	21	WC	**0.95**	**1.00**	**0.95**	
Yr 2 AP	Teacher B	21	SG	0.81	0.95	0.48	

Boldface indicates the larger mean score within each matched set.
*Teacher inadvertently skipped these questions; they were not discussed during whole class discussion.

The student written and drawn answers were scored for evidence of use of each of the two key features and the relationship between them. Even though statistical analysis of these data was not attempted, an apparent trend can be seen: In every instance where the teacher led whole class discussion about Question #7, the students in the whole class discussion format appeared to outperform the students in the small group discussion format—in each and every category. They showed more evidence for using the GPE reference line in their reasoning, more evidence for using the semi-quantitative relationships between different kinds of energy in their reasoning, and more evidence for using the relationship between the position of the GPE reference line and the amounts of *PE* and *TE*. (Because the same data was scored along all three dimensions, the results are not added across the dimensions.)

The most striking result is that *not a single student in the HP or CP small groups showed **any** evidence in **any** category*: there was no evidence for their having used either of the key features or the relationship between them in written and drawn responses on their activity sheets.

In Table 57 above, it was seen that for the simple Projectile Animations paired with activity sheets that explicitly asked students the meaning of key features, the activity sheets did not reveal a difference between the whole class and small group conditions. However, in the gravitational potential energy classes, which used a highly interactive simulation, activity sheet analysis suggests that the less advanced physics students in the study may have had trouble finding and using important features on their own. This activity sheet result is supported by videotape data from the single CP small group observed (Tables 58 and 59). The fault could have rested with the activity sheet used during Year 1; however, these teachers were very experienced and spent more time designing the sheets than a teacher would normally be likely to spend, including consulting sample activity sheets provided on the PhET simulation website and consulting with the research team. These results indicate a trend where, at least in the Gravitational Potential Energy classes and with this activity sheet, less advanced students in the whole class condition appeared much more likely to identify, and make use of the concepts supported by, key features (GPE reference line and energy bar graph) when working on their activity sheets, even though they did not have the benefit of having their hands on the computer controls.

The videotape analysis of student use of the key features produced the only result in this study that appeared to suggest an advantage for small group students, and then only for the more advanced HP and AP students. The results for the Gravitational PE videotape analysis, the Gravitational PE activity sheet analysis, and the Projectile activity sheet analysis, when taken together, show mixed results for the use of key features by higher level students but are consistent in suggesting a whole class discussion advantage for the lower level students in the four college preparatory classes in the study.

F. List of Findings

Research Question #1: *Are students able to exhibit gains in conceptual reasoning on pre-post tests from lessons that incorporate certain interactive simulations and animations?*

- All but one of the 19 class sections in the study exhibited significant gains on the short answer pre-post questions, with effect sizes ranging from small to large.

- Statistical analysis of pre-post gains on the short answer conceptual transfer questions revealed no evidence for a pre-post advantage for the small group hands-on class sections within each matched set. If anything, there appeared to be a slight trend in favor of the whole class discussion sections, especially in the Gravitational Potential Energy lesson sequence. In one comparison, the difference in gains reached significance in favor of the WC condition.

- For pre-post gains on the explanation questions, no trend was observed for the Gravitational Potential Energy lesson sequence but a trend was observed for the Projectile Motion sequence; in all five Projectile Motion comparisons, students in the whole class sections had larger mean gains.

Research Question #2: *To what extent do students and teachers engage in discussion about key concepts while working with the simulations and animations?*

- Discussion about the concepts identified as key ranged from 0% to 16% of the discussion times, from 0 to 4 ½ minutes.

- In 6 of 9 comparisons, more time was spent in the whole class discussions than in the matched small group discussions on concepts that had been identified as key. One small group did not mention the key concepts at all.

- Even when adjustment was made for the fact that the whole class discussions tended to continue longer, in 6 of 9 comparisons, the *percentage* of discussion time spent on the key concepts was greater in the whole class discussions than in the matched small group discussions. Again, there is no evidence of an advantage for the students in the small group discussions; if anything, the trend is in favor of the whole class discussions.

Research Question #3: *To what extent do teachers and students respond to conceptual difficulties and misconceptions exhibited during work with the simulations and animations?*

- From 0% to 55% of the discussion times were spent addressing student difficulties, ranging from 0 to almost 16 minutes.

- The whole class discussions spent longer responding to student difficulties in five out of the nine comparisons.

- The small group discussions spent as long or longer in four of the comparisons—but one of these small groups had the teacher present for half of the discussion time, almost the entire time difficulties were being addressed.

- Though mixed, the evidence does not support an advantage for small groups in getting conceptual difficulties of individual students addressed by discussion.

Research Question #4: *To what extent do teachers and students support the recognition, use, and interpretations of key visual features of the simulations and animations?*

- Rates of visual support episodes ranged from 0 to 207 per hour. Total numbers of episodes ranged from 0 to 54 episodes per discussion.

- For the Gravitational Potential energy sequence, in no comparison did the small group discussion show an advantage, either in rate or in total number of support episodes.

- For the Projectile Motion sequence, in three of five comparisons the whole class discussion had a greater rate of support episodes, and in four of five, the whole class discussion had a greater total number of episodes. Importantly, in the single small group that had more support episodes, *all but 6 of the 23 episodes were teacher episodes that occurred while the teacher was with the small group.* Thus, there appeared to be a slight trend in favor of the whole class discussions and a stronger trend *in favor of discussion in which a teacher was present.*

Research Question #5: *Do students recognize and use key visual features of the simulations and animations?*

- Videotape analysis of the Gravitational Potential Energy classes revealed that students in each of the lessons did recognize the features. However, analysis identified more evidence for student recognition of the features in the small groups, the only result in this study that appeared to suggest an advantage for small group students, and then only for the more advanced students. (Note that the camera was in the back of the room in the whole class discussions.)

- Activity sheet analyses of selected Projectile Motion activity sheet questions, conducted blind to condition, yielded mixed results among the matched sets; there was no clear trend in favor of either condition; a majority of students in all sections appeared able to recognize and use the key features.

- Activity sheet analyses of selected Gravitational Potential Energy questions, analyzed along three dimensions (use of concepts targeted by the two features and the relationship between them) and conducted blind to condition, yielded two striking results:

    - in every instance where the teacher facilitated whole class discussion about the questions (did not inadvertently skip them), the students in that whole class discussion outperformed the students in the matched small group discussion format—along each and every dimension;

    - the only small group students who showed evidence on their activity sheets for having used the key features were Advanced Placement students. Not a single student in the Honors Physics or College Preparatory small groups showed *any* evidence along *any* of the dimensions—they showed *no* evidence for having used either of the key features or the relationship between them in their written and drawn responses on the activity sheets.

G. Discussion of Results

1. Can the Mixed Methods Results from Chapters VI and VII Shed Light on the Pre-Post Results from Chapter V?

The teachers were surprised that, although almost all of the classes showed significant gains on the pre-post short answer questions, there appeared to be no pre-post advantage for students in the small group lesson sequences, even though small group participants had had the advantage of hands-on experience with the simulations, the time for each and every student to raise questions, a more relaxed atmosphere in which shyer students could speak up, and the engagement produced by small group work. The results of videotape analysis appear consistent with the surprising quantitative pre-post results, in that, for the most part, the videotape coding suggests no advantage for students in the small group sequences. Beyond this, the videotape and activity sheet analyses suggest several possible explanations for the quantitative results. Even though the small groups had the advantage of hands-on experience with the simulations, more time was spent on certain key concepts in the whole class discussions in 6 of 9 of the matched sets of classes. The small group discussions spent no more time than did the whole class discussions in addressing student difficulties, and in some instances, student conceptual questions were ignored or given only very brief attention in the small groups. The whole class discussions tended to have a greater number of episodes where a teacher or student provided support for using the visual features of the simulations. In the small groups, even student support episodes appeared to cluster around the teacher visits to the group. And finally, although students were observed using the visual features in small groups, not a single small group student in either of the lower level physics sections showed any evidence for having utilized these features in their thinking on the activity sheets. These results, taken together, suggest that the small group students, and especially the lower level students, may not have been able to utilize either the tools or the concepts of the simulations and animations enough during the course of their hands-on activities to confer an advantage for them over students who had experienced the simulations exclusively in the context of whole class discussion with no opportunity for hands on experience.

2. How do The Results of this Study Relate to the Literature?

These results support prior findings that novices can have numerous difficulties with interpreting visual representations whether those representations are static or dynamic (e.g., Lowe, 1995, 2003, 2004). However, the findings do not support recommendations from the literature to use simulations exclusively in a hands-on fashion. Although common sense and the literature have argued that using simulations in small groups allows students to talk more, interact with the simulation more, and become more engaged with hands-on involvement and exploration, this study exposes some disadvantages to deploying simulations exclusively in the small group setting.

The report of Jones, et al. (2001) about using animations and simulations in the chemistry classroom concludes that we know very little about how to use these materials effectively in instruction. The surprise the experienced teachers in this study exhibited concerning how students responded to the simulations in the two conditions, and the failures of both teachers and researchers to anticipate a number of student difficulties, would suggest that the Jones statement may still hold. The current results also support Cook's (2006) conclusions concerning general principles for the design of visual representations for science education. His review, which included a section on issues concerning animations in particular, concluded that design principles suggested by cognitive load theory and by laboratory work would appear to need further validation in science classroom contexts. Simulations that had been thoroughly tested by the developers with students did not always appear, within the natural classroom situations investigated here, to elicit the amount of engaged exploration these teachers had expected from their small group students. The observations of both small group and whole class discussions lend support to seminal findings of Driver (1983) and Lowe (2004) that connections obvious to teachers do not always appear to be obvious to students. Students in the present study were frequently observed missing important relationships, whether they were relationships between static visual features in simple annotated diagrams or dynamic relationships between multiple visual aids in sophisticated simulations.

Research indicates that it is important to allow students to control the pace of animation (Hegarty, Kriz, & Cate, 2003; Mayer & Chandler, 2001; Schwan & Riempp, 2004; Zahn, Barquero, & Schwan, 2004). Expert educators such as those who contributed to the report of Jones, et al. (2001) have interpreted these results to mean that hands-on use of animations is highly preferable to any other use. This interpretation of the research results is not supported by the present study. It certainly makes sense that students will benefit from being able to watch an animation or simulation at a speed that matches their comprehension of the topic. However, at times it appeared to this observer that teachers gauged this speed at least as well as students did on their own, frequently slowing down, pausing, and replaying the animations and simulations. Furthermore, when any single student in whole class called out a request for a sequence to be replayed, every student had the opportunity to witness the re-playing. In small group, students sometimes did not take the time to replay, focused on getting their activity sheets completed.

Hake's (1998) often-cited meta-study of six thousand introductory physics students persuasively showed the advantage of interactive engagement over passive methods of instruction, but looking more closely at that study offers a possible explanation consistent with the present results. Hake defines interactive engagement methods as "*those designed at least in part to promote conceptual understanding through interactive engagement of students in* ***heads-on (always) and hands-on (usually)*** *activities which yield immediate feedback through discussion with peers and/or instructors…*" (italics in original; boldface is mine). The rich whole class discussions in the current study certainly gave the appearance of "heads-on" engagement by many of the students involved. Conversely, students using "data-collection" strategies in small groups did not always appear to be engaged in a *heads-on* manner even thought they were engaged in *hands-on* experience.

The vigorous discussions in the whole class mode, together with the fact that the whole class sections performed at least as well as the small group sections, can be considered in light of the results of Hogan, Nastasi, & Pressley (2000) who found that teachers acted as catalysts in discussion, prompting students to expand and clarify their thinking. The present results are also consistent with recent findings of Smetana & Bell (2009), who compared the use of computer simulations during a chemistry unit taught by a single teacher to two class sections, one in a small group and the other in a whole class setting (N = 39). Their study found more frequent and meaningful teacher-student interactions and more frequent highly collaborative talk in the whole class group. They also found no significant difference in the pre-post gains of the two groups. The present study pursues several aspects of teacher and student interactions in detail, finding related results across two years, several teachers and levels of physics, and a number of classrooms.

3. Hypotheses and Further Discussion

Why did the hands-on small groups not do better than the whole class students? What strengths and weaknesses of the two lesson formats were suggested by videotape analysis? And lastly, what implications, if any, do these findings have for the design of instructional physics simulations?

Videotape and case study analyses suggest several grounded hypotheses for why the small groups did not do better than they did: the presence in some groups of a "get and report data" mindset, a tendency in some groups to cut off conceptual discussion in the interests of time, and a lesser amount of student- and teacher-initiated visual support episodes. Analytical codes that were used to identify elements of discussion deemed important for student learning appeared to cluster around segments in which the teacher was present. But this raises the question of *why* the small groups exhibited these characteristics. Going beyond systematically observed patterns in the data, in this section I will form some hypotheses that, if true, could help explain the findings.

For instance, something that could help explain the clustering of codes is the way in which students responded to their own conceptual difficulties. When small group students experienced a difficulty interfering with their progress through the activity sheet, they, unsurprisingly, tended to *call the teacher over*. Thus, it makes sense that both conceptual discussion and visual support would occur more frequently when the teacher was present. This does not appear to explain, however, why there tended to be more episodes of *mutual visual support* when the teacher was present. One possibility is that the teacher's presence may have helped focus the student discussion, including the student-student exchanges.

The teachers had anticipated that students would be more reluctant to raise their conceptual difficulties in whole class than in small group discussion. It was true that some whole class students were observed murmuring their dissatisfaction or puzzlement to other students rather than raising their hands to ask questions. However, one teacher developed an interesting way to deal with this; she appeared to watch for such exchanges and then to repeat the murmured comment loudly and enthusiastically to the whole class. This appeared to validate the topic as worthy of discussion, and her whole classes did, frequently, respond to this strategy with animated and engaged discourse.

There also appeared to be a difference in the ways in which whole class and small group discussions responded to unexpected time pressure. It was my impression that the small groups had a tendency to respond to such pressure by knuckling down to the task at hand, trying to complete the activity sheets. One of the videotapes clearly shows students monitoring their time and cutting short conceptual discussion in order to maintain their progress through the sheets. During whole class discussions, on the other hand, the teachers felt free, even impelled, to diverge from their lesson plans and expand the time on task when conceptual difficulties arose, even if this meant doubling the time (as in one of the analyzed discussions) or abandoning their plans for equivalent time on task (which occasionally necessitated class sections being dropped from the study). This was not only an impression gained from the videotapes but was confirmed by the teachers in follow-up interviews. Incidentally, participation in the study did not appear to be a primary cause for students' over-dedication to timely completion of the tasks. In the clearest videotaped example of this, the teacher told the students that they had an entire additional period if they needed it. In another class, students appeared to pace themselves so they could finish early and turn to another assignment, apparently from an unrelated class. It was my impression that many of these students viewed conceptual difficulties as potential derailments for their activities rather than as opportunities for learning. On the other hand, responding to the conceptual difficulties of a classmate during whole class discussion, or when the teacher was present with a small group, could give the responder a chance to shine.

The teachers reported that they tried to head off misconceptions in the whole class setting before they arose, but wanted to leave the small group students to their own devices so that the students could explore. This was not necessarily a bad thing; students certainly seemed more engaged in the small groups and reported that they liked that format better. The argument here is not that the small group work did not have benefits—it clearly did; small group students performed almost as well, if not as well, on the pre-posts as did the whole class students. Rather, the argument is that the whole class and small group formats appeared to have different strengths and weaknesses when it came to learning from the physics simulations and animations that were used.

This leads into more speculative discussion, about what other explanations might be able to account for the results. For instance, there is the possibility that the activity sheets functioned as a mediating factor, constraining the discussion and exploration to the extent that the format of the discussion was largely irrelevant. It is true that the activity sheets served as a guide to the discussions in both formats. However, they functioned differently in the two contexts. The teachers, who were accustomed to using simulations exclusively in the small group format, designed the worksheets with the small groups in mind. This was because, as one teacher explained, if an instruction on the worksheet was not relevant to the whole class situation, she could always tell students to ignore it. In addition, the teachers were accustomed to creating worksheets for their small groups but not for their whole class discussions. One reason the teachers thought at the time that the whole class discussions had not gone well is because, as they reported in follow-up interviews, they felt constrained by the worksheets in whole class discussion. All three teachers said that they felt "hamstrung" in whole class discussion, prevented from responding as spontaneously to student questions and difficulties as they normally would, because they felt obligated to follow the activity sheet. All of this would appear to have worked against the whole class students. Either the extra constraints on the whole class discussion caused by the activity sheets, and the consequent extra stress experienced by the teachers, actually helped the whole class discussions, or (more likely) there were enough strengths present in the whole class format to offset the constraining presence of the activity sheets.

Another question that arises is whether the activity sheets were simply poorly designed. As mentioned above, these experienced teachers spent more time on these activity sheets than they normally did, sending the sheets around for feedback from each other, from this researcher, and from a research advisor. The teacher who designed the activity sheet for gravitational potential energy consulted a number of the lesson plans provided on the PhET website. This teacher's first idea was to present the students with a single challenge and then to allow them open exploration for the entire two days of the lesson sequence. However, she began to become concerned about aspects of the topic she was afraid the small group students would miss and began adding in more and more specific questions and instructions. After not getting to the challenge in either class the first year, the second year she left it off the sheet altogether. A similar procedure occurred with the projectile motion activity sheets. The teachers continually discussed how they wanted to make these sheets less "cook book" and to provoke more exploration, but fears about what the small group students might miss led them inevitably to retain much of the data gathering requirements on the sheet. In the whole class situations, however, the teachers seemed to feel free to skip questions on the sheet in order to follow student suggestions, even when the suggestions were off topic or sounded silly. ("Throw the skater around!")

A possibility is that, even though the whole class and small group students within each matched set had comparable amounts of time available for learning, the time at which the post-tests were given may have disproportionately disadvantaged the small group students. The teachers were encouraged to follow the small group discussions with a whole class wrap-up in order to address any lingering questions students may have had and to make sure students had become sufficiently acquainted with the concepts addressed by the lesson. However, the teachers almost never got to the wrap-up, preferring not to interrupt small group students once they were engaged. Near the end of the study, I observed a teacher giving a wrap-up on the day following a lesson sequence. In response to a question about this during a follow-up interview, he said that he sometimes did give wrap-up discussions on a following day. It would be interesting to know whether giving the post-tests 15 minutes into the following period (for both conditions) would have yielded different comparative results. This does not explain the comparative performances on the activity sheets or the videotape coding results, however, which would not have been affected by wrap-up discussions conducted at a later time.

Another possibility is that delayed post-tests could have yielded different results. It is possible that the small group students, having worked hands-on, had more opportunity to commit what they learned to long-term memory and that the whole class students were more likely to answer from a rote repetition of what they had just heard in class. On the other hand, some students in small group appeared to come away with incorrect conclusions and it is possible that these conclusions are what would have been committed to long-term memory. Non-systematic evidence from exploratory interviews and follow-up student interviews on projectile motion revealed that some students from both discussion formats came away with misconceptions about projectile motion and that they continued to report these misconceptions several weeks after the classes. Although constraints of the present study did not allow it, it would certainly have been of interest to investigate whether small group students remembered more, and if so, what it was that they remembered several months after the lesson.

Going beyond formal evidence, my own impression is that when small groups worked well, they worked very well, but that this was a minority of the groups. In general, the whole class discussions appeared to be of a richer quality. Some students in both situations appeared to experience a lack of engagement; however, small group students who were not engaged appeared generally to remain disengaged for the whole class, while even the most disengaged whole class students appeared to become engaged from time to time, especially when the teacher did something unexpected with the simulation. In Teacher B's gravitational potential classes, especially, both in the lower-level and higher-level sections, everyone appeared to get caught up in voting on what would happen next and in calling out whimsical suggestions for what to do with the skater in the Energy Skate Park simulation.

Teacher B's discussion-leading ability leads to another observation. Although not all of the teachers in this study were as skilled as she at whole class discussion leading, they were all experienced at teaching and appeared to have a considerable amount of pedagogical content knowledge. Also, my impression was that Teachers A and B, who participated both years of the study, showed marked improvement in their discussion leading the second year they used each simulation or animation. Successfully using a simulation in the whole class condition may depend on a teacher's familiarity with the simulation as well as a willingness to change plans and follow student requests for manipulating the simulation. Also, discussions in which the teacher adopted a playful approach to the simulation or animation appeared to elicit more engagement and richer discussion than discussions approached from a more business-like goal-oriented approach.[12] Widespread successful use of simulations and animations in classrooms, no matter how carefully the tools are designed and tested, may depend on better teacher

[12] Incidentally, the second year teachers appeared to have improved in conducting small group work as well. They reported having identified provocative student questions during whole class discussions, which they then posed to individual small groups to help stimulate and focus discussion.

development, including—perhaps surprisingly—development in discussion-leading strategies and skills.

H. Limitations and Delimitations of the Study

Conclusions that can be drawn from comparisons of matched class sections in the present study are somewhat limited by the fact that the condition to which a student was assigned was determined by the class section in which the student was enrolled. The author had no control over which students were placed in which section; the participants were not randomized. Rather, sets of class sections were selected for case study comparison according to whether they were matched along several parameters (had the same teacher, the same general level of preparation, similar student demographics); once matched class sections had been identified, sections from each set were assigned a condition for the lesson sequence. The N for each comparison was limited by the number of students in available matching class sections; the result is a small N for each comparison. Therefore, the results of the quantitative comparisons cannot be projected rigorously to a population outside the study. However, this was not the intention of these comparisons; the quantitative results are intended only 1) to add quantitative detail to the individual case studies, and 2) to suggest interesting presence or absence of trends that may be worth investigating in future studies with larger samples.

The scope of the qualitative study was delimited somewhat by 1) the choice of lesson sequences to be observed and 2) the use of theoretical sampling to guide the identification of videotape and transcript segments and portions of student work for in-depth analysis, as discussed in Chapter IV Section F. As videotape analysis is a very time- and labor-intensive endeavor, only selected portions from each sequence were subjected to detailed analysis applying the full sets of codes that were developed. Also, as with any research, qualitative research involves researcher judgment in interpreting observations and in choosing for which attributes to code.

I. Implications and Suggestions for Future Research

This analysis of the use of selected animations and interactive simulations does not appear to yield much evidence for any overall advantage for small groups with hands-on access to computers over whole class discussion in which the visuals are projected onto a screen in front of the class. If anything, a slight trend was observed in favor of the whole class discussions, especially for students in the less advanced classes. These results appear to offer encouragement to teachers who do not have the resources to allow their classes to engage regularly in small group work at the computer. The whole class discussions analyzed here indicate that there appear to exist teaching strategies for promoting at least some of the active thinking and exploration that has been considered to be the strength of small group work. Furthermore, these examples suggest the somewhat surprising possibility that there may be certain instructional situations where there is an advantage to spending at least part of the time with the simulation or animation in a whole class discussion mode, for instance, to provide consistent support for students' interpretation and use of information from the visual elements on the screen.

The slight trends observed suggest that research on larger populations might yield more significant results as regards an overall advantage for a particular discussion format. However, a more productive line of research in a larger population might be to investigate in which situations one or other of the discussion formats can provide more consistent support for students and what mix of the two formats might be optimal. A single instance of a combined format occurred unexpectedly in the present study, when a teacher followed small group work with an extended whole class discussion. That class section exhibited no pre-post advantage over two matched sections in which the same teacher conducted the lesson sequence exclusively in the whole class format. It could be interesting to compare this lesson plan with the reverse order, in which a teacher engages a class in whole class discussion over a simulation and then issues a challenge for the same students to address within small groups. However, such a lesson format would appear to require a simulation rich enough to be capable of producing a suitable challenge. No doubt creative teachers could come up with other "Whole Class then Small Group" versus "Small Group then Whole Class" lesson designs. In any case, a fair comparison might be facilitated by giving the post-test the following day after whatever follow-up a teacher would normally do. It would be useful also to administer a delayed post-test at least a couple of months later.

The results of the present study suggest several design considerations for educational physics simulations and further considerations for the design process itself. Clement (1985) and Clement, Mokros, & Schultz (1986) noted a tendency for subjects to view the temporal direction on a graph as a spatial direction. In this study, the small groups in the projectile motion classes exhibited a tendency toward the reverse, *to view a spatial direction on a graph as a temporal direction*. This tendency was noted in the classes of three teachers at two different schools. In retrospect, one teacher noted that until this lesson, the students had worked only with x-t graphs. They had worked with motion maps but those maps had not had grid lines. This is one of the instances that suggested that many issues that can interfere with students' visual interpretations may not easily be anticipated by researchers or by their teachers. In particular:

- The meaning of axes may need to be given more visual support than is common (via visual cues such as quasi-realistic images, or making any rotation away from the expected orientation of the axes very explicit).

More broadly:

- Teachers may need more guidance provided along with simulations to help them identify what features and relationships are likely to be overlooked by students; teachers may also need suggestions for making these features explicit;

- Design principles for educational animations and simulations may provide guidance in design to only a first approximation;

- Successful design probably requires iterative cycles of testing and refinement and *at least some of this testing may need to be done in the noisy environment of the classroom.*

Although many physics simulations undergo iterative cycles of testing and refinement, they might benefit from trials in situations beyond the one-on-one and small group trials often employed. The fact that teachers may use simulations in the whole class situation is one reason for this suggestion. Trialing a sophisticated simulation in a whole class condition and observing teacher strategies is likely, at the very least, to yield data on teacher interactions with the tool, and may in addition suggest support materials that are needed for the teacher to make more productive use of the simulation. Such observation, along with follow-up interviews, may also suggest additional interactive features to facilitate whole class use. For instance, general design requests by teachers in the present study (requested for more than one simulation) include:

- Give the ability to mask arbitrary parts of a display so that teachers can set up novel scenarios and ask students to predict what will happen next;

- Give the ability to save multiple starting conditions so that teachers can create and test set-ups beforehand and easily switch set-ups as desired.

Trials of simulations in the context of whole class discussions may also yield data in the area of student difficulties and student-student solutions; students may observe and respond to the difficulties of their fellow classmates as much as, or even more than, in small group discussions. Although similar data can and should continue to be obtained from individual and small group trials of simulations, trials in the context of whole class discussion may be able to complement and extend these data in unexpected and productive ways.

APPENDICES

APPENDIX A

QUESTIONS FOR EXPLORATORY INTERVIEWS ON PROJECTILE MOTION (SAMPLE)

The following are about half of the questions that were used to guide (but not constrain) exploratory student interviews concerning the projectile motion trial lessons described in Chapter III. These are the questions most relevant to the larger study (Chapters IV-VIII), which was motivated and informed by these and other preliminary results discussed in Chapter III.

- Can you say anything about what worked for you or didn't work for you?
- Why do you think that helped/didn't help?
- What do you think might have helped more?
- Now can you say something about what didn't/did work for you?
- Can you say more about that?
- In general, what kinds of activities do you learn from the best?
- Do you have any suggestions about what could be improved about these lessons? If you were a researcher, what would you invent to help someone who learns things the way you do?
- I'd like to ask you some questions about the simulation in particular. What do you remember about it?
- Can you draw it?
- What can you tell me about the horizontal component of the velocity? The vertical component?

APPENDIX B

GRAVITATIONAL POTENTIAL ENERGY ACTIVITY SHEET (SAMPLE)

Name: ______________________ Date:______________

Period: _____________________ Partners: _____________________________

Online Simulation Lab → PHET: Energy Skate Park

Purpose: *The purpose of this simulation lab is to strengthen your understanding of energy conservation in real-world applications.*

Exploration Activities

Open up the University of Colorado, PhET Energy Skate Park simulation:

- Go to http://phet.colorado.edu/simulations/sims.php?sim=Energy_Skate_Park
- Click **RUN NOW**!
- Spend FIVE MINUTES to explore the simulation and familiarize yourself with the controls.

RESET Instructions: When directed to **RESET** click the reset button in the top –right corner. Then place the potential energy reference line at bottom of the track by clicking **POTENTIAL ENERGY REFERENCE** and dragging the line. This way, the skater's gravitational potential energy will be zero at the bottom.

- **RESET** and begin the exploration below, **using just the simple track provided** (you'll have a chance to do more later!).

1. Does the skater hit the same height on the opposite sides of the track?
   a. What must be true about the system for this to be possible?
   b. Click the Track Friction button to adjust the coefficient of friction. What do you observe about the skater as you adjust the setting?

Hint: **SHOW GRID** may help!

2. **RESET**. Before continuing, discuss with your partners where the gravitational potential energy, PE, the kinetic energy, KE, and the total energy the skater will be the most: at the top or bottom of the path. Fill in the prediction column in the chart below.

	PREDICT WHERE?	Try it! **ACTUALLY WHERE?**	**AMOUNT** (Round to **nearest 10 J**)
PE MOST?			
KE the MOST?			
Total Energy MOST?			

Name: ______________________ Date:_______________

3. Check your predictions by clicking on the **SHOW PATH** button and letting the rider lay down several rounds of purple dots before clicking **PAUSE**. Now click on a purple dot at the top and at the bottom to display data that will help you to determine where each type of energy is the most. (*You may need to hide the PE reference line in order to click on dots underneath it.*) Record the values in the table above, **rounding to the nearest 10 Joules**.
   a. What does total energy mean?

   b. Does KE = $\frac{1}{2}$ m v^2? (show your calculation)

   c. Does PE = mgh? (show your calculation)

Click **CHOOSE SKATER** to find its mass

Hint: You may need to move things around to see everything.

RESET. Turn on the energy Pie Chart and Bar Graph.

4. Without changing anything else, use the **CHANGE SKATER** button explore how skater's mass affects each type of energy. How does changing the skater's mass affect each type of energy?

   Potential Energy:

   Kinetic Energy:

   Total Energy:

5. Double check that the energy reference line is located a the bottom of the track, then hide it. Pick your favorite skater and use the purple dots to fill in the chart below for a position near the **TOP** of the track. Then use these values to predict the values at the **BOTTOM** of the track.
6. Check your predictions using the purple dot data and fill in the actual below:

(Round to the nearest 10 J)	**Top** (J)	**Bottom** (J)	
	Actual	Prediction	Actual
Height			
Potential Energy			
Kinetic Energy			
Speed			

Name: ______________________ Date:______________

7. Could the total energy be zero at some position? Explain.

RESET. Turn on the energy Pie Chart and Bar Graph.

8. Turn on a moderate amount of **TRACK FRICTION.**
   What happens to the maximum values of the 4 different types of energy over time?

Gravitational Potential Energy	
Kinetic Energy	
Thermal Energy	
Total Energy	

9. Turn off **TRACK FRICTION**. **PREDICT** what you think would happen to the maximum value of each type of energy if you moved the skater to Jupiter or the Moon.

PREDICTIONS	Jupiter	Moon
Gravitational PE		
Kinetic Energy		
Total Energy		

10. Try it! Record what happens and explain each case, paying particular attention to the changes in the maximum values.

Actual	Jupiter	Moon
Gravitational PE		
Kinetic Energy		
Total Energy		

Name: ______________________ Date:_______________

11. **RESET**, pull down the **TRACKS Menu** and select **LOOP**, and **SHOW GRID**. Observe what is happening and then **PAUSE** the skater, set him to start at a height of 5.5 meters, just a SMALL AMOUNT MORE than the loop height. Before hitting **RESUME**, predict whether the skater will make it all the way around the loop. .
    a. **PREDICTION** (and reasoning): .

    b. ACTUAL (and new reasoning, if you were wrong!) .

12. **RESET. Turn on the energy Pie Chart and Bar Graph.** Now select SPACE as your location. Play with the "thrusters." Describe what happens to each type of energy each time you apply the thruster rockets and explain as clearly as you can why this happens:

    Kinetic Energy:

    Potential Energy:

    Thermal Energy:

    Total Energy:

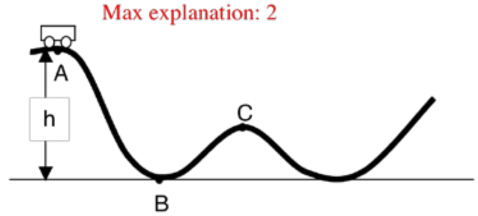
APPENDIX C

GRAVITATIONAL POTENTIAL ENERGY PRE-POST TEST (SAMPLE)

Teacher B Pre-Test Name________________________ Date___________

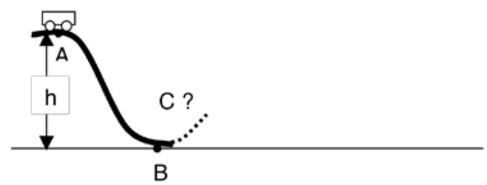
Max short answer: 9 Max explanation: 2

1. Consider the roller coaster to the right. Assume no friction or air resistance. The cart starts from rest at point A and begins to roll to the right.



a. Will the roller coaster pass over point C? Explain why or why not.
N/A (Pretest ceiling effect)

b. How high could point C be and still have the roller coaster pass over it? Sketch that here:
N/A (ambiguous)



c. Fill in the table with the cart's gravitational potential energy and kinetic energy, assuming the heights of the 3 positions are as indicated.

Each cell: 1/2 pt

	A = 10 meters	B = 0 meters	C = 5 meters
Gravitational Potential Energy	1000 J	0	500J
Kinetic Energy	0 J	1000J	500J

d. How would these values change if there were friction present? How would the values change if the roller coaster were located on the Moon? Fill in the chart below indicating INCREASE (I), DECREASE (D) or SAME (S) as compared to part (c).

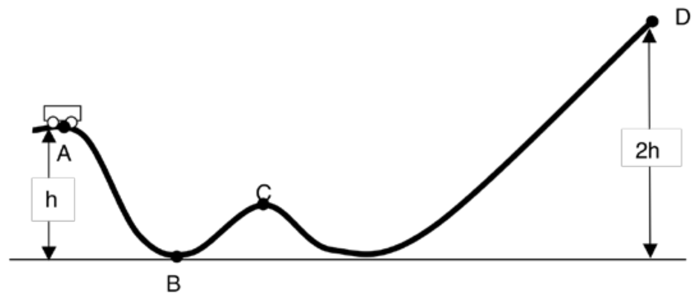
Each cell: 1/2 pt

		A	B	C
With Friction	Gravitational Potential Energy	same	same	same
	Kinetic Energy	(0)	decrease	decrease
Moon, no Friction	Gravitational Potential Energy	decrease	same	decrease
	Kinetic Energy	(0)	decrease	decrease

AP Physics 1 W10

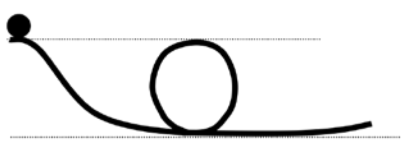
2. Back on Earth, consider the frictionless roller coaster shown below. In terms of m, g, and h (the mass of the cart, the acceleration due to gravity near the surface of Earth, and the height of Point A), how fast would the roller coaster have to be going at its start (Point A) in order to make it up to Point D?

1 Speed at A: ________√(2gh _______



3. A loop-the-loop marble track is shown below. Would a marble starting from rest and rolling to the right to make it to the end of the track. (Ignore friction and air resistance.)

*

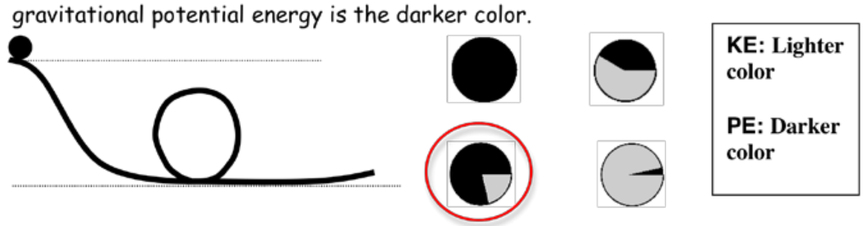


Circle: Yes / No
If No Explain:

(explanation: No, it wouldn't acquire enough KE, enough centripetal acceleration to counter gravity. Needs starting ht. of 5/4 h.)

4. For the track below, circle the pie chart that best represents the marble's energy at the top of the loop. Kinetic energy is the lighter color, gravitational potential energy is the darker color.

1



Rubric for Explanation Question 3:

0	Incorrect (or blank)
1/2	Partially correct (amount of "height" or "speed," or "too fast & will leave track")
1	Correct (amount of "potential energy" or "kinetic energy")

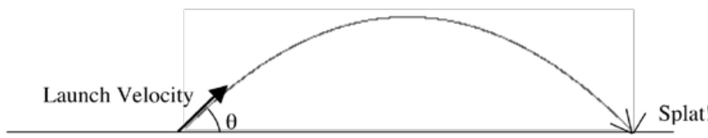
APPENDIX D

PROJECTILE MOTION PREDICTION SHEET (SAMPLE)

Projectile Motion Simulation Predictions

Key Terms:
ALTITUDE: how high the projectile goes.
RANGE – how far the projectile goes.

A projectile is launched at an angle, θ, to the horizon and lands some distance away, as shown below. Ignore the effect of air resistance.



1. What do you predict will happen to the range if you increase the launch velocity without changing the launch angle or projectile's mass?
   1) The range will increase
   2) The range will decrease
   3) The range will stay the same
   4) Other (explain...)

   Briefly explain your reasoning:

2. What do you predict will happen to the range if you shoot a more massive projectile at the same launch angle with the same speed?
   1) The range will increase
   2) The range will decrease
   3) The range will stay the same
   4) Other (explain...)

Briefly explain your reasoning:

3. For a given launch speed, what launch angle do you predict will give you the largest maximum altitude?
Briefly explain your reasoning:

4. For a given launch speed, what launch angle do you predict will give you the longest range (horizontal distance)?
Briefly explain your reasoning:

APPENDIX E

PROJECTILE MOTION DAY 1 ACTIVITY SHEET (SAMPLE)

Projectile Motion Day 1 Activity Sheet

Key Terms: ALTITUDE: how high the projectile goes. RANGE – how far the projectile goes.

Load this web application:

http://galileoandeinstein.physics.virginia.edu/more_stuff/Applets/ProjectileMotion/jarapplet.html

(Also available on your class website. To clear the screen, reload the page.)

Investigate the answers to each of the following questions. After discussion, write your answers in the spaces provided.

With air resistance off and "show trails" checked:

1. Change the launch velocity but keep all other parameters the same. Describe what happens.
2. Change the mass of the projectile but keep all the other parameters the same. Describe what happens.
3. Change the launch angle without changing the mass or the speed. Which angles give the most and least total time in the air?
4. Change the launch angle without changing the mass or the speed. Which angle gives the greatest range?

5. Keeping mass and speed the same, compare 30° and 60° launches, 40° and 50° launches. Describe the results.
6. Explain a general rule for which launch angles will result in equal ranges, assuming the mass and speed are not changed.

APPENDIX F

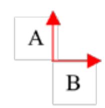
PROJECTILE MOTION DAY 2 ACTIVITY SHEET (SAMPLE)

Projectile Motion Day 2 Activity Sheet

These three movies are each a simulation of the motion of a projectile. Within the movie, there are a series of strobed snapshots that have captured the position of the ball at equal time intervals. The movie playback has been slowed down enough that we can see the motion clearly.

In each movie, we have added some red markers to call your attention to some aspect of the motion.

Quicktime Simulation #1



1) What do the red arrows indicate about the velocity?

2 a) Which component of velocity does Arrow A give you information about?

b) Is this component of the velocity changing?

3 a) Which component of velocity does Arrow B give you information about?

b) Is this component of velocity changing?

4 a) Does this simulation show acceleration? If so, in what direction?

b) What, in this simulation, lets you know that?

Compare Simulations #2 and #3.

Quicktime Simulation #2

1) What does the variable spacing between the red and blue lines indicate about the velocity?

2 a) Which component of the velocity do these lines give you information about?

b) Is this component of the velocity changing?

Quicktime Simulation #3 – ProjectileLINES

3) What does the equal spacing between the red lines indicate about the velocity?

4 a) Which component of the velocity do these lines give you information about?

b) Is this component of the velocity changing?

5 a) Does this simulation show acceleration? If so, in what direction?

b) What, in this simulation, lets you know that?

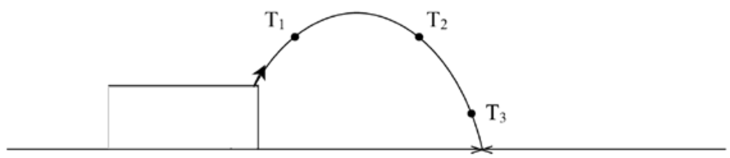
APPENDIX G

PROJECTILE MOTION PRE-POST TEST (SAMPLE)

Projectile Motion Post-Test Name: ______________________________
Teacher C Date: __________

Max multiple choice: 8 Max explanation: 2

The drawing shows the path of a baseball thrown upward at an angle from a cliff. T_1, T_2, T_3, refer to *time 1*, *time 2*, *time 3*, respectively, during the flight of the baseball. At T_1 the baseball is at the same height as it is at T_2. **IGNORE AIR RESISTANCE**.



For each question, circle the answer that agrees most closely with your thinking.

1. What do you think is happening to the vertical component of the velocity as the ball passes through each point T_1, T_2, T_3?

1 T_1: A) getting larger in upward dir. B) staying the same C) getting smaller in upward dir. D) other

1 T_2: A) getting larger in downward dir. B) staying the same C) getting smaller in downward dir. D) other

1 T_3: A) getting larger in downward dir. B) staying the same C) getting smaller in downward dir. D) other

2. What do you think is happening to the horizontal component of the velocity as it passes through each point T_1, T_2, T_3?

1 T_1: A) getting larger B) staying the same C) getting smaller D) other

1 T_2: A) getting larger B) staying the same C) getting smaller D) other

1 T_3: A) getting larger B) staying the same C) getting smaller D) other

3. If we do this with the same initial velocities on a Planet with *less* gravity than in the experiment above,

1 a) the vertical component of the velocity at time T_1 will be:

A) smaller B) larger C) same

Why?

(explanation: gravity pulling down less, reduces vertical component at a slower rate.)

1 b) the horizontal component of the velocity at time T_1 will be,

A) smaller B) larger C) same

Why?

(explanation: there is no force acting in horizontal direction in either scenario.)

S08

Rubric for Explanation Question 3a:

0	Wrong or no explanation ("it will go higher," "it will go longer")
½	Does not indicate direction of a force ("more upward velocity," "less gravity pulling")
1	Indicates presence and direction of force ("gravity pulling down less," "less vertical acceleration")

BIBLIOGRAPHY

Adams, W. K., Reid, S., LeMaster, R., McKagan, S. B., Perkins, K. K., Dubson, M., & Wieman, C. E. (2008). A study of educational simulations Part I - Engagement and learning. *Journal of Interactive Learning Research, 19*(3), 397-419.

Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology, 25*, 468-523.

Angell, C. (2004). Exploring students' intuitive ideas based on physics items in TIMSS - 1995. In C. Papanastasiou (Ed.), *The 1st IEA International Research Conference*. Lefkosia, Cyprus: International Association for the Evaluation of Educational Achievement (IEA).

Astin, A. (1993). *What matters in college: Four critical years revisited*. San Francisco: Jossey-Bass.

Bell, R., Smetana, L., & Binns, I. (2005). Simplifying Inquiry Instruction: Assessing the Inquiry Level of Classroom Activities. *The Science Teacher, 72*(7) 30-33.

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1987). Students visualization of a chemical reaction. *Education in Chemistry,* July, 117-120.

Bennett, J., Hogarth, S., Lubben, F., Campbell, B., & Robinson, A. (2010). Talking science: The research evidence on the use of small group discussions in science teaching. *International Journal of Science Education, 32* (1), 69-95.Campbell (1920)(1957)

Buckley, B. C. (2000). Interactive multimedia and model-based learning in biology. *International Journal of Science Education, 22*(9), 895–935.

Camp, C., Clement, J., Brown, D., Gonzalez, K., Kuduke, J. Minstrell, J., Schultz, K., Steinberg, M., Veneman, V., & Zietsman, A. (1994). *Preconceptions in Mechanics: Lessons Dealing With Conceptual Difficulties*. Dubuque, Iowa: Kendall Hunt.

Clement, J. (1985). Misconceptions in graphing. In the *Proceedings of the Ninth Conference of the International Group for the Psychology of Mathematics Education*, Noordwijkerhout, The Netherlands, July, 1985.

Clement, J. (1989a). The concept of variation and misconceptions in Cartesian graphing. *Focus on Learning Problems in Mathematics, 11*(2), 77-87.

Clement, J. (1989b). Learning via Model Construction and Criticism: Protocol Evidence on Sources of Creativity in Science. In J. Glover, R. Ronning, and C. Reynolds (Eds.), *Handbook of Creativity: Assessment, Theory and Research* (p. 341-381) NY: Plenum.

Clement, J. (1994). Use of physical intuition and imagistic simulation in expert problem solving. In Tirosh, D. (Eds.), *Implicit and explicit knowledge: An educational approach.* Ablex Publishing Corp.

Clement**,** J. (2000a). Analysis of clinical interviews: Foundations and model viability. In Lesh, R. & Kelly, A. (eds.), *Handbook of research methodologies for science and mathematics education* (pp. 341-385). Hillsdale, NJ: Lawrence Erlbaum.

Clement, J. (2000b). Model based learning as a key research area for science education. *International Journal of Science Education, 22*(9), pp. 1041-1053.

Clement, J. (2002). Managing student/teacher co- construction of visualizable models in large group discussion. *Proceedings of AETS 2002 National Meeting.*

Clement, J. (2006). Thought experiments and imagery in expert protocols. In L. Magnani (ed.) *Model based reasoning in science and engineering*. London: King's College Publications.

Clement, J., Mokros, J., & Schultz, K. (1986). Adolescents' graphing skills: A descriptive analysis. Paper presented at AERA. ERIC RIE #ED-264-127

Clement, J., & Rea-Ramirez, M.A. (1998). The role of dissonance in conceptual change. *Proceedings of the National Association for Research in Science Teaching Conference*.

Clement, J., & Steinberg, M. (2002). Step-wise evolution of models of electric circuits: A "learning-aloud" case study. *Journal of the Learning Sciences, 11*, 389-452.

Clement, J., Zietsman, A., & Monaghan, J. (2005). Imagery in science learning in students and experts. In J. Gilbert (Ed.), *Visualization in science education* (pp. 169-184). Dordrecht, The Netherlands: Springer.

Collins, A., & Gentner, D. (1987). How people construct mental models. In Holland, D. & Quinn, N. (Eds.). *Cultural Models in Thought & Language*. Cambridge Univ. Press.

Cook, M. (2006). Visual Representations in Science Education: The Influence of Prior Knowledge and Cognitive Load Theory on Instructional Design Principles. *Science Education 90,* 1073-1091. DOI 10.1002/sce.

Creswell, J. W. (1998). *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Sage Publications. Thousand Oaks, CA. pp. 202-203.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Darden, L. (1991). *Theory Change in Science: Strategies from Mendelian Genetics* (New York: Oxford).

DiSessa, A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction, 22*(3), 293-331.

Driver, R. (1983) The fallacy of induction in science teaching. *The Pupil as Scientist?* Chapt. 1. Milton Keynes: Open University Press. pp. 1-10.

Driver, R., & Bell, B. (1986). Students' thinking and the learning of science: A constructivist view. *School Science Review, 67*(240), 443-56.

Dwyer, F. (1987). Enhancing visualized instruction: Recommendations for practitioners. State College, PA: Learning Services.

Finke, R. (1989). *Principles of Mental Imagery* (Cambridge, MA: MIT Press), p.2.

Forbus, K., & Gentner, D. (1997). Qualitative mental models: Simulations or memories? *Proceedings of the Eleventh International Workshop on Qualitative Reasoning*, Cortona, Italy

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gentner, D. (2002). Psychology of mental models. In N. J. Smelser & P. B. Bates (Eds.), *International Encyclopedia of the Social and Behavioral Sciences* (pp. 9683-9687). Amsterdam: Elsevier Science.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner and A. L. Stevens (Eds.) *Mental Models* (Hillside, NJ: Erlbaum), p.99.

Giere, R. (1988). *Explaining Science: A Cognitive Approach* (Chicago: Chicago University Press).

Gilbert, S. W. (1999). *The model as a vehicle for understanding the nature and processes of science.* Report. http://eric.ed.gov/PDFS/ED443657.pdf

Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Chicago; Aldine.

Goldman, S. R. (2003). Learning in complex domains: When and why do multiple representations help? *Learning and Instruction 13*, 239-244.

Gooding, D. (1992). What is experimental about thought experiments? In D. Hull, M. Forbes & K. Okruhlick (Eds.), *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 2: Symposia and Invited Papers*, (pp. 280-290). East Lansing, MI: Philosophy of Science Association.

Gooding, D. (1996). Creative rationality: towards an abductive model of scientific change. *Philosophica 58* (2), pp. 73-102.

Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and Learning. In D.C. Berliner & R.C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 15-46). New York: Simon & Schuster Macmillan.

Hake, Richard R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66* (1), 64-74.

Halloun, I., & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics, 53*(11), pp. 1043-1048.

Hammer, D. (1995). Student inquiry in a physics class discussion. *Cognition and Instruction, 13*(3), 401-430.

Hammer, D. (1997). Discovery Learning and Discovery Teaching. *Cognition and Instruction, 15*(4), 485-529

Harrison, A. G., & Treagust, D.F. (2000). A typology of school science models. *International Journal of Science Education, 22*, 1011–1026.

Hatano, G. (1993). Time to merge Vygotskian and constructivist conceptions of knowledge acquisition. In E.A. Forman, N. Minick, & C.A. Stone (Eds.), *Contexts for Learning: Sociocultural Dynamics in children's Development* (pp. 153-166). New York: Oxford University Press.

Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems, *Journal of Experimental Psychology: Learning, Memory and Cognition 18*, 1084.

Hegarty, M. (2004). Dynamic visualizations and learning: getting to the difficult questions. *Learning and Instruction, 14*, 343-351.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition & Instruction, 21*, 325–360.

Hegarty, M., Narayanan, N., & Freitas, P. (2002). Understanding machines from multimedia and hypermedia presentations. In J. Otero, J. A. Leon & A. Graesser, Eds. *The Psychology of Science Text Comprehension*, pp. 357–384. Hillsdale, NJ: Lawrence Erlbaum.

Herron, M.D. (1971). The nature of scientific enquiry. *Educational Psychologist*, *79*(2): 171–212.

Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics, 55*(5), 440-454.

Hestenes, D. (1990). Secrets of genius, *New Ideas in Psychology 2*, 231.

Hestenes, D. (1996). *Modeling methodologies for physics teachers*. Paper presented at the International Conference on Undergraduate Physics Education. College Park, MD, August, 1996.

Hogan, K., & Pressley, M. (1997). Scaffolding scientific competencies within classroom communities of inquiry. In Hogan, K. & Pressley, M. (Eds.) *Scaffolding student learning : instructional approaches and issues.* Cambridge, MA Brookline Books.

Hogan, K., Nastasi, B.K., & Pressley, M. (2000). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction, 17*(4), pp. 379-432.

Inagaki, K., Morita, E., & Hatano. G. (1999). Teaching–learning of evaluative criteria for mathematical arguments through classroom discourse: a cross-national study. *Mathematical Thinking and Learning, I*(2), 93-111.

Johnson, B., & Onwugbuzie, A. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14-26.

Johnson-Laird, P. N. (1975). Models of deduction. In R. J. Falmagne (Ed.) *Reasoning: Representation and process in children and adults.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science, 4*, 71-115.

Jones, G., & Carter, C. (1997). Small groups and shared constructions, in *Teaching science for understanding: A human constructivist view*, edited by Joel J. Mintzes, James H. Wandersee, and Joseph D. Novak. (Academic Press: New York) pp. 261-279

Jones, L., Jordan, K., & Stillings, N. (2001). Molecular visualization in science education (2001). *Report from the Molecular Visualization in Science Education Workshop*, NCSA Access Center, Arlington, VA, January 12-14.

Just, M., & Carpenter, P. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual difference in spatial ability. *Psychological Review, 92,* 137-172.

Khan, S. (2003). *Model construction processes in a chemistry class*. Paper presented at the National Association of Research in Science Teaching Conference, Philadelphia, PA, March 23-26, 2003.

Khan, S. (2008a). What if scenarios for testing student models in chemistry. In J. Clement & M. A. Rea-Ramirez (Eds.), *Model based learning and instruction in science* (Ch. 8, pp. 139-150). Dordrecht: Springer

Khan, S. (2008b). Co-construction and model evolution in chemistry. In J. Clement & M. A. Rea-Ramirez (Eds.), *Model based learning and instruction in science* (Ch. 4, pp. 59-78). Dordrecht: Springer

Kintsch, W. (1986). Learning from text, *Cognition & Instruction **3***, 87.

Kintsch, W. (1988). The role of knowledge in discourse-comprehension: A construction-integration model. *Psychological Review **95***, 163.

Kozhevnikov, M., Hegarty, M., & Mayer, R. (1999). Students' use of imagery in solving qualitative problems in kinematics, in *Proceedings of the Annual Meeting of the American Educational Research Association* (AERA, Montreal, 1999).

Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory & Cognition (*Jun 2005) *33*(4), 710.

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction, 13,* 205–226.

Kozma, R., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching, 34*(9), 949-968.

Larkin, J., & Chabay, R. (1989). Research on teaching scientific thinking: Implications for computer-based instruction. In L. Resnick and L. Klopfer (Eds.), *Toward the Thinking Curriculum* (pp. 150-172). Alexandria, VA: Association for Supervision and Curriculum Development.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science, 208* (June), pp. 1335-1342.

Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science 11*, 65-99.

Linn, M. (2003). Technology and science education: Starting points, research programs, and trends. *International Journal of Science Education, 25*(6), 727–758.

Linn, M., & Peterson, A. (1985). Emergence and Characterization of Sex Differences in Spatial Ability: A Meta-Analysis. *Child Development*, 56, 6, 1479-1498.

Lowe, R. K. (1995). Selectivity in diagrams: Reading beyond the lines. *Educational Psychology, 14*(4), pp. 4667-4692.

Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction 13,* 157-176.

Lowe, R. K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction 14*(3), 257-274.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research, 59*(1), pp.43-64.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*(1), 1-19.

Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction 13*, 125-139

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: an experimental test of a dualcoding hypothesis. *Journal of Educational Psychology*, *83*, 484–490.

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology 88*(1), 64-73

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*(2), 390-397.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312-320

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review 14*(1) 87-99.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312-320.

McDermott, L., Rosenquist, M., and van Zee, E. (1986). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics 55*(6), 503-513.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

Menand, L. (1997). *Pragmatism: A reader*. New York: Vintage.

Miller, A. (1986). *Imagery in Scientific Thought* (Cambridge, MA: MIT Press).

Monaghan, J. M., & Clement, J. (1999). Use of a computer simulation to develop mental simulations for understanding relative motion concepts. *Int. J. Sci. Educ., 21*(9), 921-944.

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 9*1(2), 358-368.

Mokros, J. R., & Tinker, R. F. (1987). The impact of micro-computer based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, 24(4), 369-383.

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91*(2), 358-368.

Narayanan, N. H., & Hegarty, M. (2002). Multimedia design for communication of dynamic information. *International Journal of Human–Computer Studies, 57*, 279–315.

Nassaji, H., & Wells, G. (2000). What's the use of 'triadic dialogue'?: An investigation of teacher-student interaction. *Applied Linguistics, 21*(3), 376-406.

Nersessian, N. (1995). Should physicists preach what they practice? Constructive modeling in doing and learning physics. *Science & Education, 4*(3), 203-226.

Nersessian, N. (2007). Mental modeling in conceptual change, in S.Vosniadou (Ed.), *Handbook of Conceptual Change* (Mahwah, NJ: Lawrence Erlbaum, 2007).

Nunez-Oviedo, M. C. (2004). *Teacher-student co-construction processes in biology: Strategies for developing mental models in large group discussions*. Unpublished Doctoral Dissertation. University of Massachusetts, Amherst.

Nunez Oviedo, M. C., & Clement, J. (2003). *Model evolution: A strategy based on model based teaching and learning theory*. Presented at AERA conference, Chicago, IL, April 21-25

Nunez-Oviedo, M. C., & Clement, J. (2008). A competition strategy and other discussion modes for developing mental models in large group discussion. In J. Clement & M. A. Rea-Ramirez (Eds.), *Model based learning and instruction in science* (Ch. 7, pp. 117-138). Dordrecht: Springer.

Nussbaum, J., & Novick, S. (1982). Alternative frameworks, conceptual conflict and accommodation: Toward a principled teaching strategy. *Instructional Science, 11*(3), 183-200.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Peirce, C. S. (1878). How to make our ideas clear. *Popular Science Monthly, 12*, 286–302.

Perkins, K., Adams, W., Dubson, M., Finkelstein, N., Reid, S., Wieman, C., & LeMaster, R. (2006). PhET: Interactive simulations for teaching and learning physics. *The Physics Teacher, 44*, 18-23.

Piaget, J. (1963). *The origins of intelligence in children*. New York: W. W. Norton.

Price, N. (2007). Self-study of the evolution of a “deferred judgment questioning” discussion mode in a middle school science teacher. *Proceedings of the 2007 Annual Meeting of the National Association for Research in Science Teaching* (NARST), New Orleans, April 14-19.

Price, N., & Clement, J. (2011). Comparative case studies of discussion strategies used in dynamic computer simulation and static image-based lessons. *Proceedings of the 2011 Annual Meeting of the National Association for Research in Science Teaching* (NARST), Orlando, FL, April 3-6.

Raghavan, K., Sartoris, M., & Zimmerman, C. (2002). *Impact of model-centered instruction on student learning.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, April 7-10.

Raghavan, K., Sartoris, M., & Glaser, R. (1998). Why does it go up? The impact of the MARS curriculum as revealed through changes in student explanations of a helium balloon. *Journal of Research in Science Teaching, 35*(5), 547-567.

Rea-Ramirez, M. A. (1999). *Developing complex mental models through explanatory need.* Paper presented at NARST, March 1999, Boston, MA.

Rea-Ramirez, M. A., & Nunez-Oviedo, M. C. (2008). Role of discrepant questioning leading to model element modification. In J. Clement & M. A. Rea-Ramirez (Eds.), *Model based learning and instruction in science* (Ch. 11, pp. 195-214). Dordrecht: Springer.

Rea-Ramirez, M.A., Nunez-Oviedo, M.C., Clement, J., & Else, M. J. (2004). *Energy in the Human Body Curriculum.* University of Massachusetts, Amherst.

Redish**,** E. F. (1994). The implications of cognitive studies for teaching physics. *American Journal of Physics, 62*(6), 796-803.

Reid, D. J., Zhang, J., & Chen, Q. (2003). Supporting scientific discovery learning in a simulation environment. *Journal of Computer Assisted Learning, 19*, 9– 20.

Reid, S., Adams, W., Dubson, M., Loeblein, T., Perkins, K., & Wieman, C. (2009). *Energy Skate Park v2.05.* PhET Interactive Simulations, University of Colorado. http://phet.colorado.edu/index.php

Reiner, M., Pea, R. D., & Shulman, D. J. (1994). The impact of simulator-based instruction on diagramming in geometrical optics by introductory physics students. *Journal of Science Education and Technology*, October.

Rieber, L. (1990a). Animation in computer-based instruction. *Educational Technology Research and Development, 38*(1), 77-86.

Rieber, L. (1990b). Using computer-animated graphics in science instruction with children. *Journal of Education al Psychology, 82*(1), 135-140.

Roth, W.-M. (2001). Learning science through technological design. *Journal of Research in Science Teaching, 38*(7), 768-790.

Russell, J., & Kozma, R. (2005). Assessing learning from the use of multimedia chemical visualization software. In J. K. Gilbert (Ed.), *Visualization in science education* (pp. 229-332). Dordrecht, The Netherlands: Springer.

Russell, T., & Osborne, J. (1993). *Constructivist research, curriculum development and practice in primary classrooms: Reflections on five years of activity in the Science Processes and Concept Exploration (SPACE) Project.* Paper presented at the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics, Cornell, University, NY. July, 1993)

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction, 13*(2), 141-156.

Schwan, S., & Riempp, R. (2004) The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction 14*(3), 293-305.

Seddon, G., Eniaiyeju, P., & Jusoh, I. (1984). The visualization of rotation in diagrams of three dimensional structures. *American Educational Research Journal*, 21 (1), 25-38.

Seddon, G., & Tariq, R. (1982). The visualization of spatial transformations in diagrams of molecular structures. *European Journal of Science Education*, 4, 409-420

Sellares, J., & Toussaint, G. (2003). On the role of kinesthetic thinking in computational geometry, *International Journal of Mathematical Education in Science and Technology* ***34***, 219.

Smetana, L., & Bell, R. (2009). Incorporation of computer simulations in whole-class vs. small-group settings. *Proceedings of the 2009 Annual Meeting of the National Association for Research in Science Teaching* (NARST), Garden Grove, CA.

Smith, C., Maclin, D., Grosslight, L., & Davis, H. (1997). Teaching for understanding: A study of students' preinstruction theories of matter and a comparison of the effectiveness of two approaches to teaching about matter and density. *Cognition and Instruction, 15*(3), 317-393.

Smith, K. A. (1996). Cooperative learning: Making "groupwork" work. In T. E. Sutherland, C. C. Bonwell (Eds.), *Using active learning in college classes: A range of options for faculty*. San Francisco: Jossey-Bass.

Steinberg, M. S. (2004). *The CASTLE project student manual*. Pasco Scientific.

Stephens, L., & Clement, J. (2012). "The role of thought experiments in science and science learning," In B. Fraser, K. Tobin, & C. McRobbie (Eds.), *Second International Handbook of Science Education: Springer International Handbooks of Education, Vol 24*, Part 2 (pp. 157-175). Dordrecht: Springer. DOI: 10.1007/978-1-4020-9041-7_13

Stephens, L., & Clement, J. (2010). "Documenting the use of expert scientific reasoning processes by high school physics students," *Physical Review Special Topics – Physics Education Research, **6***, 020122. http://link.aps.org/doi/10.1103/PhysRevSTPER.6.020122

Stephens, L. & Clement, J. (2009). "Extreme case reasoning and model based learning in experts and students", *Proceedings of the 2009 Annual Meeting of the National Association for Research in Science Teaching* (NARST), Garden Grove, CA.

Stephens, L., & Clement, J. (2006). Depictive gestures as evidence for dynamic mental imagery in four types of student reasoning. In L. McCullough, L. Hsu, & P. Heron (Eds.), *2006 Physics Education Research Conference: AIP Conference Proceedings*, Vol. 883, p.89. Melville, NY: American Institute of Physics.

Strauss, A., & Corbin, J. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Sage Publications. Thousand Oaks, CA., pp. 12 & 101.

Thompson, Bruce (1999). Statistical significance tests, effect size reporting and the vain pursuit of pseudo-objectivity. *Theory Psychology*; 9; 191.

Trickett, S. B., & Trafton, J. G. (2002). The instantiation and use of conceptual simulations in evaluating hypotheses: movies in the mind in scientific reasoning. *Proceedings of the 24th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

Tversky, B., Bauer-Morrisson, J., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies. 57*(4), 247-262.

van Zee, E., & Minstrell, J. (1997a). Reflective discourse: Developing shared understandings in a physics classroom. *International Journal of Science Education, 19*, pp. 209-228.

van Zee, E., & Minstrell, J. (1997b). Using questioning to guide student thinking. *The Journal of the Learning Sciences, 6*(2), pp. 227-269.

von Glasersfeld, E. (1989). Cognition, construction of knowledge, and teaching. *Synthese, 80* (1), 121-140.

Vosniadou, S. (2002). Mental models in conceptual development. In Magnani, L. & N. Nersessian (Eds.) *Model-Based Reasoning: Science, Technology, Values*. (pp. 353-368). New York, Plenum Publishers.

Vygotsky, L. S. (1978). *Mind and society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wells, M., Hestenes, D., & Swackhamer, S. (1995). A modeling method for high school physics instruction, *Am. J. Phys. 63*(7), 606-619.

Williams, E. G., & Clement, J. (2007). Identifying Model-Based Teaching Strategies: A Case Study of Two High School Physics Teachers. *Proceedings of the NARST Annual Meeting* – New Orleans, LA, April, 2007.

Williams, E. G., & Clement, J. (2009). Model co-construction in high school physics: A case study of teachers' intended instructional pathways and recovery routes. *Proceedings of the NARST Annual Meeting* – Garden Grove, CA, April, 2009

Williams, E. G., & Clement, J. (2010). Supporting Students' Construction of Mental Models for Electric Circuits: An Investigation of Teacher Moves Used in Whole Class Discussions. *Proceedings of the NARST Annual Meeting* – Philadelphia, PA, March, 2010.

Williams, M., Linn, C., Ammon, P., & Gearhart, M. (2004). Learning to teach inquiry science in a technology-based environment: A case study. *Journal of Science Education and Technology, 13*(2), 189-206.

Wood, D. J., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychiatry and Psychology, 17*(2), 89-100.

Woods, D., & Fassnacht, C. (2007). Transana v2.30b. http://www.transana.org. Madison, WI: The Board of Regents of the University of Wisconsin System.

Zahn, C., Barquero, B., & Schwan, D. (2004). Learning with hyperlinked videos—design criteria and efficient strategies for using audiovisual hypermedia. *Learning and Instruction 14*(3), 275-291.

Zietsman, A. I., & Hewson, P. W. (1986). Effect of instruction using microcomputer simulations and conceptual change strategies on science learning. *Journal of Research in Science Teaching, 23*(1), 27-39.